PCT

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : **C07K 14/415, C12N 15/29, 15/82**	**A2**	(11) International Publication Number: **WO 98/04586** (43) International Publication Date: 5 February 1998 (05.02.98)

(21) International Application Number: PCT/GB97/02046

(22) International Filing Date: 29 July 1997 (29.07.97)

(30) Priority Data:
9615879.5	29 July 1996 (29.07.96)	GB
9622626.1	30 October 1996 (30.10.96)	GB
9704789.8	7 March 1997 (07.03.97)	GB

(71) Applicant (for all designated States except US): JOHN INNES CENTRE INNOVATIONS LIMITED [GB/GB]; Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): SCHULZE-LEFERT, Paul, Maria, Josef [DE/GB]; Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB). PANSTRUGA, Ralph [DE/GB]; Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB). BÜSCHGES, Rainer [DE/GB]; Sainsbury Laboratory, Norwich Research Park, Colney Lane, Norwich NR4 7UH (GB).

(74) Agents: WALTON, Seán, M. et al.; Mewburn Ellis, York House, 23 Kingsway, London WC2B 6HP (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: POLYNUCLEOTIDE AND ITS USE FOR MODULATING A DEFENCE RESPONSE IN PLANTS

(57) Abstract

The cloning of the *Mlo* gene of Barley, *mlo* mutants, and homologues from other species including rice and *Arabidopsis thaliana* enables manipulation of pathogen defence responses in plants. Nucleic acid and polypeptides are used in the production of transgenic plants in which the pathogen defence response is modulated, particularly stimulated. Various approaches enable diagnostic determination of the presence of susceptibility or resistance alleles in plants. Compounds able to modulate a defence response by interaction with Mlo or mlo protein are identifiable using assays involving the protein or fragments thereof.

POLYNUCLEOTIDE AND ITS USE FOR MODULATING A DEFENCE RESPONSE IN PLANTS

The present invention relates to stimulating a defence
response in plants, with a view to providing the plants with

5   enhanced pathogen resistance. More specifically, it has
resulted from cloning of the barley *Mlo* gene, various mutant
mlo alleles, and a number of homologues from various species.
The *Mlo* gene has been isolated using a positional cloning
approach which has never previously been successful in Barley.

10   Details and discussion are provided below. Wild-type *Mlo*
exerts a negative regulatory function on a pathogen defence
response, such that mutants exhibit a defence response in the
absence of pathogen. In accordance with the present invention,
down-regulation or out-competition of *Mlo* function may be used

15   to stimulate a defence response in transgenic plants,
conferring increased pathogen resistance.

Mutations have been described in several plants in which
defence responses to pathogens appear to be constitutively
expressed. Mutation-induced recessive alleles (*mlo*) of the

20   barley *Mlo* locus exhibit a leaf lesion phenotype and confer an
apparently durable, broad spectrum resistance to the powdery
mildew pathogen, *Erysiphe graminis* f sp *hordei*.

Resistance responses to the powdery mildew pathogen have
been genetically well characterized (Wiberg, 1974; Søgaard and

25   Jørgensen, 1988; Jørgensen, 1994). In most analyzed cases
resistance is specified by race-specific resistance genes
following the rules of Flor's gene-for-gene hypothesis (Flor,
1971). In this type of plant/pathogen interaction, resistance

2

is specified by and dependent on the presence of two
complementary genes, one from the host and one from the fungal
pathogen. The complementary genes have been termed
operationally (pathogen) resistance ("R") gene and avirulence
5   gene, respectively. Most of the powdery mildew resistance genes
(Mlx) act as dominant or semidominant traits (Jørgensen, 1994).

Monogenic resistance mediated by recessive (mlo) alleles
of the Mlo locus is different. Apart from being recessive, it
differs from race-specific resistance to single pathogen
10   strains in that (i) it confers broad spectrum resistance to
almost all known isolates of the pathogen (ii) mlo resistance
alleles have been obtained by mutagen treatment of any tested
susceptible wild type (Mlo) variety, and (iii) mlo resistance
alleles exhibit a defence mimic phenotype in the absence of the
15   pathogen (Wolter et al. , 1993). Thus, the genetic data
indicate the Mlo wild type allele exerts a negative regulatory
function on defence responses to pathogen attack.

Resistance mediated by mlo alleles is currently widely
used in barley breeding and an estimated 10 million hectares
20   are annually planted in Europe with seeds of this genotype. A
'mlo like' inherited resistance to powdery mildew in other
cereal plants has not been reported so far although the fungus
is a relevant pathogen in wheat (attacked by Erysiphe graminis
f sp tritici), oat (attacked by E. g. f sp avenae), and rye
25   (attacked by E. g. f sp secalis). Because cereals are
morphologically, genetically and biochemically highly related
to each other (Moore et al., 1995), one would predict the

3

existence of homologous genes in these species. The failure to

have found a 'mlo like' inherited resistance in wheat and oat

is probably due to their hexaploid genomes, making it difficult

to obtain by mutagenesis defective alleles in all six gene

5   copies, and the chance of all such mutations occurring in

Nature is remote. The failure to have found a mlo equivalent

in other cereals is probably due to insignificant amount of

mutational analysis in these species and complications as a

result of their outbreeding nature (e.g. rye).

10   RFLP markers closely linked to Mlo on barley chromosome 4

were previously identified on the basis of a mlo backcross line

collection containing mlo alleles from six genetic backgrounds

(Hinze et al. , 1991). The map position of Mlo on the basis of

RFLP markers was consistent with its chromosomal localization

15   as determined by a previous mapping with morphological markers

(Jørgensen, 1977).

Having identified an ~3cM genetic interval containing Mlo

bordered by genetic markers, we decided to attempt to isolate

the gene via positional cloning.

20   However, there is no documented example of a successful

positional cloning attempt of a barley gene. We were faced

with a number of difficulties.

Firstly, the genome of barley (5.3×10^9 bp/haploid genome

equivalent; Bennett and Smith, 1991) has almost double the size

25   of the human genome and because the total genetic map covers

~1.800 cM (Becker et al., 1995) we were confronted with a very

unfavourable ratio of genetic and physical distances (1 cM

4

corresponds to ~ 3 Mb).

Secondly, a high resolution genetic map had to be constructed around *Mlo* enabling the positioning of linked markers with a precision of better than 0.1 cM.

5   Thirdly, we aimed to physically delimit the target gene and both flanking DNA markers on individual large insert genomic clones, a procedure later termed "chromosome landing" (Tanksley *et al.*, 1995). For this purpose, a complete barley YAC library from barley Megabase DNA had to be constructed with
10   an average insert size of 500-600 kb, which was unprecedented.

Fourthly, we had to prepare unusual genetic tools that enabled us to identify the *Mlo* gene within a physically delimited region without the need for a time consuming generation of barley transgenic plants and testing of different
15   candidate genes. We used for our studies ten characterized radiation- or chemically-induced *mlo* mutants (Jørgensen, 1992). For a conclusive chain of evidence of the gene isolation we decided to depend upon a functional restoration of the wild type *Mlo* allele starting out from characterized *mlo* defective
20   alleles. For this purpose, we performed *mlo* heteroallelic crosses and isolated susceptible intragenic *Mlo* recombinants. The sequence analysis of these proves the function of the described gene.

The cloning of the barley *Mlo* gene and homologues,
25   including homologues from other plant species, gives rise to a number of practical applications, reflected in the various aspects of the present invention.

5

According to a first aspect of the present invention there
is provided a nucleic acid molecule comprising a nucleotide
sequence encoding a peptide with Mlo function. Those skilled in
the art will appreciate that "Mlo function" refers to the
ability to suppress a defence response, said defence response
being race and/or pathogen independent and autonomous of the
presence of a pathogen, such as, for example, the *Mlo* gene of
barley, the *Acd* gene and the *Lsd* gene of *Arabidopsis*.

mlo mutations that down-regulate or disrupt functional
expression of the wild-type Mlo sequence are recessive, such
that they are complemented by expression of a wild-type
sequence. Thus "*Mlo* function" can be determined by assessing
the level of constitutive defence response and/or
susceptibility of the plant to a pathogen such as, for example,
powdery mildew or rust (e.g. yellow rust). Accordingly, a
putative nucleotide sequence with *Mlo* function can be tested
upon complementation of a suitable *mlo* mutant. The term "mlo
function" is used to refer to sequences which confer a *mlo*
mutant phenotype on a plant.

The capitalisation of "Mlo" and non-capitalisation of
"mlo" is thus used to differentiate between "wild-type" and
"mutant" function.

A *mlo* mutant phenotype is characterised by the exhibition
of an increased resistance against one or more pathogens, which
is race and/or pathogen independent and autonomous of the
presence of a pathogen.

The test plant may be monocotyledonous or dicotyledonous.

6

Suitable monocots include any of barley, rice, wheat, maize or
oat, particularly barley. Suitable dicots include *Arabidopsis*.

Nucleic acid according to the invention may encode a
polypeptide comprising the amino acid sequence shown in Figure
2, or an allele, variant, derivative or mutant, or homologue,
thereof.

Nucleic acid according to the present invention may have
the sequence of a *Mlo* gene of barley, or be a mutant, variant
(or derivative) or allele of the sequence provided, or a
homologue thereof. Preferred mutants, variants and alleles are
those which encode a sequence which retains a functional
characteristic of the wild-type gene, especially the ability to
suppress a defence response as discussed herein. Other
preferred mutants, variants and alleles encode a sequence
which, in a homozygote, cause constitutive activation of a
defence response, or at least promotes activation of a defence
response (i.e. is a *mlo* mutant sequence), e.g. by reducing or
wholly or partly abolishing Mlo function. Preferred mutations
giving *mlo* mutant sequences are shown in Table 1. Changes to a
sequence, to produce a mutant, derivative or variant, may be by
one or more of addition, insertion, deletion or substitution of
one or more nucleotides in the nucleic acid, leading to the
addition, insertion, deletion and/or substitution of one or
more amino acids. Of course, changes to the nucleic acid which
make no difference to the encoded amino acid sequence are
included. Particular variants, mutants, alleles and

7

derivatives are discussed further below, as well as homologues.

A preferred nucleic acid sequence according to an aspect of the present invention is shown in Figure 2 along with the predicted amino acid sequence. Nucleic acid may be subject to

5   alteration by way of substitution of nucleotides and/or a combination of addition, insertion and/or substitution of one or more nucleotides with or without altering the encoded amino acids sequence (by virtue of the degeneracy of the genetic code).

10   As discussed below, further aspects of the present invention provide homologues of the *Mlo* sequence shown in Figure 2, including from rice (genomic sequence Figure 5, bottom line, cDNA sequence Figure 10, amino acid sequence Figure 13) and barley (genomic sequence Figure 6, bottom line,

15   cDNA sequence Figure 11, amino acid sequence Figure 14); also Table 5B (nucleotide sequences) and Figure 5A (amino acid sequences) show homologous EST's from rice and *Arabidopsis*.

The present invention also provides a vector which comprises nucleic acid with any one of the provided sequences,

20   preferably a vector from which a product can be expressed. The vector is preferably suitable for transformation into a plant cell and/or a microbial cell. The invention further encompasses a host cell transformed with such a vector, especially a plant cell or a microbial cell (e.g. *Agrobacterium tumefaciens*).

25   Thus, a host cell, such as a plant cell, comprising nucleic acid according to the present invention is provided. Within the cell, the nucleic acid may be incorporated within the nuclear

8

genome, i.e. a chromosome. There may be more than one
heterologous nucleotide sequence per haploid genome.

A vector comprising nucleic acid according to the present
invention need not include a promoter, particularly if the

5   vector is to be used to introduce the nucleic acid into cells
for recombination into the genome.

Nucleic acid molecules and vectors according to the
present invention may be provided in a form isolated and/or
purified from their natural environment, in substantially pure

10   or homogeneous form, or free or substantially free of nucleic
acid or genes of the species of interest or origin other than
the relevant sequence. Nucleic acid according to the present
invention may comprise cDNA, RNA, genomic DNA and may be wholly
or partially synthetic. The term "isolate" may encompass all

15   these possibilities.

The present invention also encompasses the expression
product of any of the nucleic acid sequences disclosed and
methods of making the expression product by expression from
encoding nucleic acid therefore under suitable conditions in

20   suitable host cells, e.g. *E. coli*. Those skilled in the art
are well able to construct vectors and design protocols for
expression and recovery of products of recombinant gene
expression. Suitable vectors can be chosen or constructed,
containing one or more appropriate regulatory sequences,

25   including promoter sequences, terminator fragments,
polyadenylation sequences, enhancer sequences, marker genes and
other sequences as appropriate. For further details see, for

example, Molecular Cloning: a Laboratory Manual: 2nd edition,

Sambrook et al, 1989, Cold Spring Harbor Laboratory Press.

Transformation procedures depend on the host used, but are well

known. Many known techniques and protocols for manipulation of

5   nucleic acid, for example in preparation of nucleic acid

constructs, mutagenesis, sequencing, introduction of DNA into

cells and gene expression, and analysis of proteins, are

described in detail in *Short Protocols in Molecular Biology*,

Second Edition, Ausubel et al. eds., John Wiley & Sons, 1992.

10   The disclosures of Sambrook et al. and Ausubel et al. are

incorporated herein by reference, along with all other

documents mentioned.

      Purified Mlo protein, or a fragment, mutant or variant

thereof, e.g. produced recombinantly by expression from

15   encoding nucleic acid therefor, may be used to raise antibodies

employing techniques which are standard in the art. Antibodies

and polypeptides comprising antigen-binding fragments of

antibodies may be used in identifying homologues from other

species as discussed further below.

20   Methods of producing antibodies include immunising a

mammal (eg human, mouse, rat, rabbit, horse, goat, sheep or

monkey) with the protein or a fragment thereof. Antibodies may

be obtained from immunised animals using any of a variety of

techniques known in the art, and might be screened, preferably

25   using binding of antibody to antigen of interest. For

instance, Western blotting techniques or immunoprecipitation

may be used (Armitage et al, 1992, Nature 357: 80-82).

10

Antibodies may be polyclonal or monoclonal.

As an alternative or supplement to immunising a mammal, antibodies with appropriate binding specificity may be obtained from a recombinantly produced library of expressed

5   immunoglobulin variable domains, eg using lambda bacteriophage or filamentous bacteriophage which display functional immunoglobulin binding domains on their surfaces; for instance see WO92/01047.

Antibodies raised to a polypeptide or peptide can be used

10   in the identification and/or isolation of homologous polypeptides, and then the encoding genes. Thus, the present invention provides a method of identifying or isolating a polypeptide with Mlo or mlo function (in accordance with embodiments disclosed herein), comprising screening candidate

15   peptides or polypeptides with a polypeptide comprising the antigen-binding domain of an antibody (for example whole antibody or a fragment thereof) which is able to bind an Mlo or mlo peptide, polypeptide or fragment, variant or variant thereof or preferably has binding specificity for such a

20   peptide or polypeptide, such as having an amino acid sequence identified herein. Specific binding members such as antibodies and polypeptides comprising antigen binding domains of antibodies that bind and are preferably specific for a Mlo or mlo peptide or polypeptide or mutant, variant or derivative

25   thereof represent further aspects of the present invention, as do their use and methods which employ them.

Candidate peptides or polypeptides for screening may for

11

instance be the products of an expression library created using

nucleic acid derived from an plant of interest, or may be the

product of a purification process from a natural source.

A peptide or polypeptide found to bind the antibody may be

5   isolated and then may be subject to amino acid sequencing. Any

suitable technique may be used to sequence the peptide or

polypeptide either wholly or partially (for instance a fragment

of a polypeptide may be sequenced). Amino acid sequence

information may be used in obtaining nucleic acid encoding the

10   peptide or polypeptide, for instance by designing one or more

oligonucleotides (e.g. a degenerate pool of oligonucleotides)

for use as probes or primers in hybridisation to candidate

nucleic acid, or by searching computer sequence databases, as

discussed further below.

15   A further aspect of the present invention provides a

method of identifying and cloning Mlo homologues from plants,

including species other than Barley, which method employs a

nucleotide sequence derived from that shown in Figure 2.

Further similar aspects employ a nucleotide sequence derived

20   from any of the other Figures provided herein. Nucleic acid

libraries may be screened using techniques well known to those

skilled in the art and homologous sequences thereby identified

then tested. The provision of sequence information for the Mlo

gene of Barley and various homologues enables the obtention of

25   homologous sequences from Barley and other plant species, as

exemplified further herein.

Also, one can easily derive PCR primers based on putative

12

exon sequences, which might be identified by comparison with

the *Mlo* sequence provided in Figure 2 wherein exons are

highlighted, and perform RT-PCR with total RNA from the plant

of interest, e.g. barley and rice for the homologues shown in

5   Figures 5 and 6, with cDNA and amino acid sequences shown in

other figures herein.

The homologues whose nucleotide sequences are given and

whose amino acid sequences are given or are deducible represent

and provide further aspects of the present invention in

10   accordance with those disclosed for the Barley gene shown in

Figure 2.

The present invention also extends to nucleic acid

encoding a *Mlo* homologue obtained using a nucleotide sequence

derived from that shown in Figure 2, or the amino acid sequence

15   shown in Figure 2. Preferably, the nucleotide sequence and/or

amino acid sequence shares homology with the sequence encoded

by the nucleotide sequence of Figure 2, preferably at least

about 50%, or at least about 55%, or at least about 60%, or at

least about 65%, or at least about 70%, or at least about 75%,

20   or at least about 80% homology, or at least about 85% homology,

or at least about 90% homology, most preferably at least about

95% homology. "Homology" in relation to an amino acid sequence

may be used to refer to identity or similarity, preferably

identity. High levels of amino acid identity may be limited to

25   functionally significant domains or regions.

A mutant, allele, variant or derivative amino acid

sequence in accordance with the present invention may include

13

within the sequence shown in Figure 2, a single amino acid

change with respect to the sequence shown in Figure 2, or 2, 3,

4, 5, 6, 7, 8, or 9 changes, about 10, 15, 20, 30, 40 or 50

changes, or greater than about 50, 60, 70, 80 or 90 changes.

5   In addition to one or more changes within the amino acid

sequence shown in Figure 2, a mutant, allele, variant or

derivative amino acid sequence may include additional amino

acids at the C-terminus and/or N-terminus.

As is well-understood, homology at the amino acid level is

10   generally in terms of amino acid similarity or identity.

Similarity allows for "conservative variation", i.e.

substitution of one hydrophobic residue such as isoleucine,

valine, leucine or methionine for another, or the substitution

of one polar residue for another, such as arginine for lysine,

15   glutamic for aspartic acid, or glutamine for asparagine.

Similarity may be as defined and determined by the TBLASTN

program, of Altschul et al. (1990) J. Mol. Biol. 215: 403-10,

which is in standard use in the art, or, and this may be

preferred, the standard program BestFit, which is part of the

20   Wisconsin Package, Version 8, September 1994, (Genetics

Computer Group, 575 Science Drive, Madison, Wisconsin, USA,

Wisconsin 53711). BestFit makes an optimal alignment of the

best segment of similarity between two sequences. Optimal

alignments are found by inserting gaps to maximize the number

25   of matches using the local homology algorithm of Smith and

Waterman

Homology may be over the full-length of the relevant

14

sequence shown herein, or may more preferably be over a
contiguous sequence of about or greater than about 20, 25, 30,
33, 40, 50, 67, 133, 167, 200, 233, 267, 300, 333, 400, 450,
500, 550, 600 or more amino acids or codons, compared with the
5   relevant amino acid sequence or nucleotide sequence as the case
may be.

The EST sequences provided herein, have on average 70%
similarity and 50% identity with the Mlo amino acid sequence of
Figure 2. We show that the rice homologue (Figure 5) and
10 barley homologue (Figure 6) have an amino acid identity of 81%
(amino acid sequences shown in Figure 13 and Figure 14).

In certain embodiments, an allele, variant, derivative,
mutant or homologue of the specific sequence may show little
overall homology, say about 20%, or about 25%, or about 30%, or
15 about 35%, or about 40% or about 45%, with the specific
sequence. However, in functionally significant domains or
regions the amino acid homology may be much higher. Putative
functionally significant domains or regions can be identified
using processes of bioinformatics, including comparison of the
20 sequences of homologues. Functionally significant domains or
regions of different polypeptides may be combined for
expression from encoding nucleic acid as a fusion protein. For
example, particularly advantageous or desirable properties of
different homologues may be combined in a hybrid protein, such
25 that the resultant expression product, with Mlo or mlo
function, may comprise fragments of various parent proteins.

The nucleotide sequence information provided herein, or

15

any part thereof, may be used in a data-base search to find

homologous sequences, expression products of which can be

tested for *Mlo* or *mlo* function. These may have ability to

complement a *mlo* mutant phenotype in a plant or may, upon

5   expression in a plant, confer a *mlo* phenotype.

In public sequence databases we recently identified

several homologues for the sequence of Figure 2. We have

already found homologues in rice and barley, and the dicot.

Arabidopsis.

10   By sequencing homologues, studying their expression

patterns and examining the effect of altering their expression,

genes carrying out a similar function to *Mlo in* Barley are

obtainable. Of course, mutants, variants and alleles of these

sequences are included within the scope of the present

15   invention in the same terms as discussed above for the Barley

gene.

Homology between the homologues as disclosed herein, may

be exploited in the identification of further homologues, for

example using oligonucleotides (e.g. a degenerate pool)

20   designed on the basis of sequence conservation.

According to a further aspect, the present invention

provides a method of identifying or a method of cloning a *Mlo*

homologue, e.g. from a species other than Barley, the method

employing a nucleotide sequence derived from that shown in

25   Figure 2 or that shown in any of the other Figures herein. For

instance, such a method may employ an oligonucleotide or

oligonucleotides which comprises or comprise a sequence or

16

sequences that are conserved between the sequences of Figures 2

and/or 5 and/or 6 and/or 10 and/or 11 and/or 12, or encoding an

amino acid sequence conserved between Figure 2 and/or 7 and/or

13 and/or 14 and/or 15 to search for homologues. Thus, a

5   method of obtaining nucleic acid is provided, comprising

hybridisation of an oligonucleotide or a nucleic acid molecule

comprising such an oligonucleotide to target/candidate nucleic

acid. Target or candidate nucleic acid may, for example,

comprise a genomic or cDNA library obtainable from an organism

10   known to contain or suspected of containing such nucleic acid,

either monocotyledonous or dicotyledonous. Successful

hybridisation may be identified and target/candidate nucleic

acid isolated for further investigation and/or use.

Hybridisation may involve probing nucleic acid and

15   identifying positive hybridisation under suitably stringent

conditions (in accordance with known techniques) and/or use of

oligonucleotides as primers in a method of nucleic acid

amplification, such as PCR. For probing, preferred conditions

are those which are stringent enough for there to be a simple

20   pattern with a small number of hybridisations identified as

positive which can be investigated further. It is well known

in the art to increase stringency of hybridisation gradually

until only a few positive clones remain.

As an alternative to probing, though still employing

25   nucleic acid hybridisation, oligonucleotides designed to

amplify DNA sequences may be used in PCR reactions or other

methods involving amplification of nucleic acid, using routine

17

procedures. See for instance "PCR protocols; A Guide to Methods and Applications", Eds. Innis *et al*, 1990, Academic Press, New York.

Preferred amino acid sequences suitable for use in the design of probes or PCR primers for some purposes are sequences conserved (completely, substantially or partly) between at least two Mlo peptides or polypeptides encoded by genes able to suppress a defence response in a plant, e.g. with any of the amino acid sequences of any of the various figures herein and/or encoded by the nucleotide sequences of any of the various figures herein.

On the basis of amino acid sequence information oligonucleotide probes or primers may be designed, taking into account the degeneracy of the genetic code, and, where appropriate, codon usage of the organism from the candidate nucleic acid is derived.

Preferably an oligonucleotide in accordance with certain embodiments of the invention, e.g. for use in nucleic acid amplification, is up to about 50 nucleotides, or about 40 nucleotides or about 30 or fewer nucleotides in length (e.g. 18, 21 or 24).

Assessment of whether or not such a PCR product corresponds to *Mlo* homologue genes may be conducted in various ways. A PCR band from such a reaction might contain a complex mix of products. Individual products may be cloned and each one individually screened. It may be analysed by transformation to assess function on introduction into a plant

of interest.

As noted, nucleic acid according to the present invention
is obtainable using oligonucleotides, designed on the basis of

5   sequence information provided herein, as probes or primers.
Nucleic acid isolated and/or purified from one or more cells of
barley or another plant (see above), or a nucleic acid library
derived from nucleic acid isolated and/or purified from the
plant (e.g. a cDNA library derived from mRNA isolated from the

10   plant), may be probed under conditions for selective
hybridisation and/or subjected to a specific nucleic acid
amplification reaction such as the polymerase chain reaction
(PCR). The nucleic acid probed or used as template in the
amplification reaction may be genomic DNA, cDNA or RNA. If

15   necessary, one or more gene fragments may be ligated to
generate a full-length coding sequence.

We have tested several PCR primers derived from the *Mlo*
sequence disclosed herein to test their specificity for
amplifying nucleic acid according to the present invention,

20   using both barley genomic DNA and RT-PCR templates. The latter
was synthesized from barley polyA$^+$ RNA. In each case we were
able to amplify the expected *Mlo* derived gene fragments as
shown by cloning and subsequent DNA sequencing of the PCR
products. Full length cDNA clones can be obtained as described

25   by 5' and 3' RACE technology if RT-PCR products are used as
templates.

Examples of primers tested include:

19

25L	5'-GTG CAT CTG CGT GTG CGT A-3'
25LN	5'-GTG TGC GTA CCT GGT AGA G-3'
25R	5'-AAC GAC GTC TGG TGC GTG-3'
33	5'-TGC AGC TAT ATG ACC TTC CCC CTC-3'
37	5'-GGA CAT GCT GAT GGC TCA GA-3'
38	5'-CAG AAC TTG TCT CAT CCC TG-3
38A	5'-GGC TAT ACA TTG GGA CTA ACA-3'
38B	5'-CGA ATC ATC ACA TCC TAT GTT-3'
39	5'-GCA AGT TCG ACT TCC AC-3'
39A	5'-TCG ACT TCC ACA AGT ACA TCA-3'
53	5'-AGC GTA CCT GCG TAC GTA G-3'

Various primer combinations have been tested:

38/39A; 38/39; 38/33; 38/37; 38A/39A; 38B/39A; 38/25L; 38/25LN; 25R/25L; 25R/25LN; 25R/53.

Various aspects of the present invention include the obtainable nucleic acid, methods of screening material, e.g. cell lysate, nucleic acid preparations, for the presence of nucleic acid of interest, methods of obtaining the nucleic acid, and the primers and primer combinations given above.

The sequence information provided herein also allows the design of diagnostic tests for determination of the presence of a specific mlo resistance allele, or a susceptibility allele (e.g. wild-type), in any given plant, cultivar, variety, population, landrace, part of a family or other selection in a breeding programme or other such genotype. A diagnostic test

20

may be based on determination of the presence or absence of a

particular allele by means of nucleic acid or polypeptide

determination.

At the nucleic acid level, this may involve hybridisation

5   of a suitable oligo- or poly-nucleotide, such as a fragment of

the Mlo gene or a homologue thereof, including any homologue

disclosed herein, or any particular allele, such as an allele

which gives an mlo phenotype, such as any such allele disclosed

herein. The hybridisation may involve PCR designed to amplify

10   a product from a given allelic version of mlo, with subsequent

detection of an amplified product by any of a number of

possible methods including but not limited to gel

electrophoresis, capillary electrophoresis, direct

hybridisation of nucleotide sequence probes and so on. A

15   diagnostic test may be based on PCR designed to amplify various

alleles or any allele from the Mlo locus, with a test to

distinguish the different possible alleles by any of a number

of possible methods, including DNA fragment size, restriction

site variation (e.g. CAPS - cleaved amplified polymorphic

20   sites) and so on.   A diagnostic test may also be based on a

great number of possible variants of nucleic acid analysis that

will be apparent to those skilled in the art, such as use of a

synthetic mlo-derived sequence as a hybridisation probe.

Broadly, the methods divide into those screening for the

25   presence of nucleic acid sequences and those that rely on

detecting the presence or absence of a polypeptide. The

methods may make use of biological samples from one or more

21

plants or cells that are suspected to contain the nucleic acid

sequences or polypeptide.

Exemplary approaches for detecting nucleic acid or

polypeptides include analysing a sample from the plant or plant

5   cell by:

(a) comparing the sequence of nucleic acid in the sample

with all or part of the nucleotide sequence shown in Figure 7

to determine whether the sample from the patient contains a

mutation;

10   (b) determining the presence in the sample of a

polypeptide including the amino acid sequence shown in Figure 2

or a fragment thereof and, if present, determining whether the

polypeptide is full length, and/or is mutated, and/or is

expressed at the normal level;

15   (c) performing DNA fingerprinting to compare the

restriction pattern produced when a restriction enzyme cuts

nucleic acid in the sample with the restriction pattern

obtained from the nucleotide sequence shown in Figure 7 or from

a known mutant, allele or variant thereof;

20   (d) contacting the sample with a specific binding member

capable of binding to nucleic acid including the nucleotide

sequence as set out in Figure 7 or a fragment thereof, or a

mutant, allele or variant thereof, the specific binding member

including nucleic acid hybridisable with the sequence of Figure

25   7 or a polypeptide including a binding domain with specificity

for nucleic acid including the sequence of Figure 7 or the

polypeptide encoded by it, or a mutated form thereof, and

22

determining binding of the specific binding member;

(e) performing PCR involving one or more primers based on the nucleotide sequence shown in Figure 7 to screen the sample for nucleic acid including the nucleotide sequence of Figure 7

5   or a mutant, allele or variant thereof.

When screening for a resistance allele nucleic acid, the nucleic acid in the sample will initially be amplified, e.g. using PCR, to increase the amount of the analyte as compared to other sequences present in the sample. This allows the target

10   sequences to be detected with a high degree of sensitivity if they are present in the sample. This initial step may be avoided by using highly sensitive array techniques that are becoming increasingly important in the art.

A variant form of the gene may contain one or more

15   insertions, deletions, substitutions and/or additions of one or more nucleotides compared with the wild-type sequence (such as shown in Table 1) which may or may not disrupt the gene function. Differences at the nucleic acid level are not necessarily reflected by a difference in the amino acid

20   sequence of the encoded polypeptide. However, a mutation or other difference in a gene may result in a frame-shift or stop codon, which could seriously affect the nature of the polypeptide produced (if any), or a point mutation or gross mutational change to the encoded polypeptide, including

25   insertion, deletion, substitution and/or addition of one or more amino acids or regions in the polypeptide. A mutation in a promoter sequence or other regulatory region may prevent or

23

reduce expression from the gene or affect the processing or stability of the mRNA transcript.

Tests may be carried out on preparations containing genomic DNA, cDNA and/or mRNA. Testing cDNA or mRNA has the advantage of the complexity of the nucleic acid being reduced by the absence of intron sequences, but the possible disadvantage of extra time and effort being required in making the preparations. RNA is more difficult to manipulate than DNA because of the wide-spread occurrence of RN'ases.

Nucleic acid in a test sample may be sequenced and the sequence compared with the sequence shown in Figure 2, or other figure herein, to determine whether or not a difference is present. If so, the difference can be compared with known susceptibility alleles (e.g. as summarised in Table 1) to determine whether the test nucleic acid contains one or more of the variations indicated, or the difference can be investigated for association with disease resistance.

The amplified nucleic acid may then be sequenced as above, and/or tested in any other way to determine the presence or absence of a particular feature. Nucleic acid for testing may be prepared from nucleic acid removed from cells or in a library using a variety of other techniques such as restriction enzyme digest and electrophoresis.

Nucleic acid may be screened using a variant- or allele-specific probe. Such a probe corresponds in sequence to a region of the gene, or its complement, containing a sequence alteration known to be associated with disease resistance.

24

Under suitably stringent conditions, specific hybridisation of such a probe to test nucleic acid is indicative of the presence of the sequence alteration in the test nucleic acid. For efficient screening purposes, more than one probe may be used

5   on the same test sample.

Allele- or variant-specific oligonucleotides may similarly be used in PCR to specifically amplify particular sequences if present in a test sample. Assessment of whether a PCR band contains a gene variant may be carried out in a number of ways

10   familiar to those skilled in the art. The PCR product may for instance be treated in a way that enables one to display the mutation or polymorphism on a denaturing polyacrylamide DNA sequencing gel, with specific bands that are linked to the gene variants being selected.

15   An alternative or supplement to looking for the presence of variant sequences in a test sample is to look for the presence of the normal sequence, e.g. using a suitably specific oligonucleotide probe or primer.

Approaches which rely on hybridisation between a probe and

20   test nucleic acid and subsequent detection of a mismatch may be employed. Under appropriate conditions (temperature, pH etc.), an oligonucleotide probe will hybridise with a sequence which is not entirely complementary. The degree of base-pairing between the two molecules will be sufficient for them to anneal

25   despite a mis-match. Various approaches are well known in the art for detecting the presence of a mis-match between two annealing nucleic acid molecules.

25

For instance, RN'ase A cleaves at the site of a mis-match.
Cleavage can be detected by electrophoresing test nucleic acid
to which the relevant probe or probe has annealed and looking
for smaller molecules (i.e. molecules with higher

5   electrophoretic mobility) than the full length probe/test
hybrid. Other approaches rely on the use of enzymes such as
resolvases or endonucleases.

Thus, an oligonucleotide probe that has the sequence of a
region of the normal gene (either sense or anti-sense strand)

10   in which mutations associated with disease resistance are known
to occur (e.g. see Table 1) may be annealed to test nucleic
acid and the presence or absence of a mis-match determined.
Detection of the presence of a mis-match may indicate the
presence in the test nucleic acid of a mutation associated with

15   disease resistance. On the other hand, an oligonucleotide
probe that has the sequence of a region of the gene including a
mutation associated with disease resistance may be annealed to
test nucleic acid and the presence or absence of a mis-match
determined. The presence of a mis-match may indicate that the

20   nucleic acid in the test sample has the normal sequence, or a
different mutant or allele sequence. In either case, a battery
of probes to different regions of the gene may be employed.

The presence of differences in sequence of nucleic acid
molecules may be detected by means of restriction enzyme

25   digestion, such as in a method of DNA fingerprinting where the
restriction pattern produced when one or more restriction
enzymes are used to cut a sample of nucleic acid is compared

26

with the pattern obtained when a sample containing the normal
gene or a variant or allele is digested with the same enzyme or
enzymes.

The presence of absence of a lesion in a promoter or other
5   regulatory sequence may also be assessed by determining the
level of mRNA production by transcription or the level of
polypeptide production by translation from the mRNA.

Nucleic acid isolated and/or purified from one or more
cells of a plant or a nucleic acid library derived from nucleic
10   acid isolated and/or purified from cells (e.g. a cDNA library
derived from mRNA isolated from the cells), may be probed under
conditions for selective hybridisation and/or subjected to a
specific nucleic acid amplification reaction such as the
polymerase chain reaction (PCR).

15   A method may include hybridisation of one or more (e.g.
two) probes or primers to target nucleic acid. Where the
nucleic acid is double-stranded DNA, hybridisation will
generally be preceded by denaturation to produce single-
stranded DNA. The hybridisation may be as part of a PCR
20   procedure, or as part of a probing procedure not involving PCR.
An example procedure would be a combination of PCR and low
stringency hybridisation. A screening procedure, chosen from
the many available to those skilled in the art, is used to
identify successful hybridisation events and isolate hybridised
25   nucleic acid.

Binding of a probe to target nucleic acid (e.g. DNA) may
be measured using any of a variety of techniques at the

27

disposal of those skilled in the art. For instance, probes may

be radioactively, fluorescently or enzymatically labelled.

Other methods not employing labelling of probe include

examination of restriction fragment length polymorphisms,

5   amplification using PCR, RNAase cleavage and allele specific

oligonucleotide probing.

Probing may employ the standard Southern blotting

technique. For instance DNA may be extracted from cells and

digested with different restriction enzymes. Restriction

10   fragments may then be separated by electrophoresis on an

agarose gel, before denaturation and transfer to a

nitrocellulose filter. Labelled probe may be hybridised to the

DNA fragments on the filter and binding determined. DNA for

probing may be prepared from RNA preparations from cells.

15   Preliminary experiments may be performed by hybridising

under low stringency conditions various probes to Southern

blots of DNA digested with restriction enzymes. Suitable

conditions would be achieved when a large number of hybridising

fragments were obtained while the background hybridisation was

20   low. Using these conditions nucleic acid libraries, e.g. cDNA

libraries representative of expressed sequences, may be

searched.

As noted, those skilled in the art are well able to employ

suitable conditions of the desired stringency for selective

25   hybridisation, taking into account factors such as

oligonucleotide length and base composition, temperature and so

on.

28

In some preferred embodiments of diagnostic assays
according to the present invention, oligonucleotides according
to the present invention that are fragments of any of the
sequences shown in Figure 2, or any allele associated with
5   disease resistance, e.g. as identified in Table 1, are at least
about 10 nucleotides in length, more preferably at least about
15 nucleotides in length, more preferably at least about 20
nucleotides in length, more preferably about 30 nucleotides in
length. Such fragments themselves individually represent
10   aspects of the present invention. Fragments and other
oligonucleotides may be used as primers or probes as discussed
but may also be generated (e.g. by PCR) in methods concerned
with determining the presence in a test sample of a sequence
indicative of disease resistance.

15

There are various methods for determining the presence or
absence in a test sample of a particular polypeptide, such as
the polypeptide with the amino acid sequence shown in Figure 2,
or other figure herein, or an amino acid sequence mutant,
20   variant or allele thereof (e.g. including an alteration shown
in Table 1).

A sample may be tested for the presence of a binding
partner for a specific binding member such as an antibody (or
mixture of antibodies), specific for one or more particular
25   variants of the polypeptide shown in Figure 2, e.g. see Table
1.

In such cases, the sample may be tested by being contacted

29

with a specific binding member such as an antibody under

appropriate conditions for specific binding, before binding is

determined, for instance using a reporter system as discussed.

Where a panel of antibodies is used, different reporting labels

5   may be employed for each antibody so that binding of each can

be determined.

       A specific binding member such as an antibody may be used

to isolate and/or purify its binding partner polypeptide from a

test sample, to allow for sequence and/or biochemical analysis

10   of the polypeptide to determine whether it has the sequence

and/or properties of the wild-type polypeptide or a particular

mutant, variant or allele thereof. Amino acid sequence is

routine in the art using automated sequencing machines.

15   The use of diagnostic tests for mlo alleles allows the

researcher or plant breeder to establish, with full confidence

and independent from time consuming resistance tests, whether

or not a desired allele is present in the plant of interest (or

a cell thereof), whether the plant is a representative of a

20   collection of other genetically identical plants (e.g. an

inbred variety or cultivar) or one individual in a sample of

related (e.g. breeders' selection) or unrelated plants. The

mlo alleles conferring the desirable disease resistance

phenotype are recessive, and are not therefore detectable at

25   the whole plant phenotype level when in a heterozygous

condition in the presence of a wild-type Mlo allele.

Phenotypic screening for the presence of such recessive alleles

30

is therefore only possible on material homozygous for the *mlo*
locus and so delays substantially the generation in a plant
breeding programme at which selection can be reliably and cost-
effectively applied.   In a backcross breeding programme where,

5   for example, a breeder is aiming to introgress a desirable *mlo*
allele into an elite adapted high performing target genotype,
the *mlo* locus will be permanently in the heterozygous condition
until selfing is carried out.   Nucleic acid or polypeptide
testing for the presence of the recessive allele avoids the

10   need to test selfed progeny of backcross generation
individuals, thus saving considerable time and money.   In other
types of breeding scheme based on selection and selfing of
desirable individuals, nucleic acid or polypeptide diagnostics
for the desirable *mlo* alles in high throughput, low cost assays

15   as provided by this invention, reliable selection for the
desirable *mlo* alleles can be made at early generations and on
more material than would otherwise be possible.   This gain in
reliability of selection plus the time saving by being able to
test material earlier and without costly resistance phenotype

20   screening is of considerable value in plant breeding.

     By way of example for nucleic acid testing, the barley
mlo-5 resistance allele is characterized by a G- to A-
nucleotide substitution in the predicted start codon of the *Mlo*
gene (Table 1).   The mutation may easily be detected by

25   standard PCR amplification of a *Mlo* gene segment from genomic
template DNA with the primers:

          forward primer: 5'-GTTGCCACACTTTGCCACG-3'

31

reverse primer: 5'-AAGCCAAGACGACAATCAGA-3'

(for example), followed by digestion witht he restriction

enzyme *PshA1*. This generates a cleaved amplified polymorphic

sequences (CAPS) marker which may be displayed using

5 conventional agarose gel electrophoresis. Presence of a 769 bp

fragment is indicative of the presence of the *mlo-5* allele.

The *mlo-9* resistance allele is characterized by a C- to T-

nucleotide substitution (Table 1). This allele is of

particular relevance since it is used frequently in breeding

10 material. The mutational event may be easily detected using

the primers:

forward primer 5'-GRRGCCACACTTTGCCACG-3'

reverse primer 5'-AAGCCAAGACGACAATCAGA-3'

(for example) and subsequent digestion of genomic amplification

15 products with the restriction enzyme *HhaI*. This generates a

CAPS marker which may be displayed by conventional agarose gel

electrophoresis. The presence of a 374 bp fragment is

indicative of the presence of *mlo-9*.

A third, particularly interesting allele is *mlo-12*,

20 characterised by a substitution a residue 240, specifically a

Phe240 to leucine replacement. This may result from a C720 to

A substitution in the encoding nucleotide sequence (Table 1).

This is the only currently documented *mlo* allele for which

conclusive evidence is available that the altered protein

25 retains residual wild-type activity (Hentrich, 1979, *Arch.*

Züchtungsvorsch., Berlin 9, S. 283-291). *mlo-12* exhibits no

detectable spontaneous cell death reaction but confers a

32

sufficient level of resistance to pathogens such as the powdery

mildew fungus. *mlo-12* may therefore be the allele of choice in

breeding programs if minimal pleiotropic effects (spontaneous

cell death) are desirable after introgression of the *mlo*

5   resistance in elite breeding lines. Furthermore, the molecular

site of the amino acid substitution within the Mlo protein

allows the design of alleles with a residual wild-type

activity, and also the obtention of interacting and/or

inhibitory molecules, reducing undesirable pleiotropic effects

10   from a complete loss of function of the Mlo protein.

Nucleic acid-based determination of the presence or

absence of *mlo* alleles may be combined with determination of

the genotype of the flanking linked genomic DNA and other

unlinked genomic DNA using established sets of markers such as

15   RFLPs, microsatellites or SSRs, AFLPs, RAPDs etc. This enables

the researcher or plant breeder to select for not only the

presence of the desirable *mlo* allele but also for individual

plant or families of plants which have the most desirable

combinations of linked and unlinked genetic background. Such

20   recombinations of desirable material may occur only rarely

within a given segregating breeding population or backcross

progeny. Direct assay of the *mlo* locus as afforded by the

present invention allows the researcher to make a stepwise

approach to fixing (making homozygous) the desired combination

25   of flanking markers and mlo alleles, by first identifying

individuals fixed for one flanking marker and then identifying

progeny fixed on the other side of the *mlo* locus all the time

33·

knowing with confidence that the desirable *mlo* allele is still

present.

The present disclosure provides sufficient information for

a person skilled in the art to obtain genomic DNA sequence for

5   any given new or existing *mlo* allele and devise a suitable

nucleic acid- and/or polypeptide-based diagnostic assay.

Existing *mlo* alleles to which this may be applied include, for

example, *mlo-1*, *mlo-3*, *mlo-4*, *mlo-5*, *mlo-6*, *mlo-7*, *mlo-8*, *mlo-*

9, *mlo-10*, *mlo-12*, *mlo-13*, *mlo-16*, *mlo-17*, *mlo-26* and *mlo-28*,

10   for all of which sequence information is provided herein (see

e.g. Figure 2 and Table 1).   In designing a nucleic acid assay

account is taken of the distinctive variation in sequence that

characterises the particular variant allele.   Thus, the present

invention extends to an oligonucleotide fragment of a *mlo*

15   allele, having a sequence which allows it to hybridise

specifically to that allele as compared with other *mlo* alleles.

Such an oligonucleotide spans a nucleotide at which a *mlo*

mutation occurs, and may include the mutated nucleotide at or

towards its 3' or 5' end.   Such an oligonucleotide may

20   hybridise with the sense or anti-sense strand.   The variation

may be within the coding sequence of the <u>mlo</u> gene, or may lie

within an intron sequence or in an upstream or downstream non-

coding sequence, wherein disruption affects or is otherwise

related to the lesion in *Mlo* that results in the mildew

25   resistant phenotype.

The *mlo-9* allele is widely but not exclusively used in

plant breeding (J Helms Jorgensen - *Euphytica* (1992) 63: 141-

34

152), *mlo*-11 is also used. Use of *mlo* mutants in practical

breeding has largely been restricted to spring barley, because

the spontaneous cell death response associated with many of the

mutant alleles appears to represent a penalty to plant growth

5   and performance when incorporated into high yielding winter

barley genotypes. However different *mlo* alleles have

different degrees of associated spontaneous cell death

response, and thus some, either existing or newly created from

mutagenesis programmes or isolated as spontaneous mutants, are

10   more suitable than others for incorporation into winter barley

backgrounds. The *mlo*-12 allele may be particularly suitable

since no detectable pleiotropic effects occur despite

conferring a sufficient level of pathogen resistance. The use

of mlo based mildew resistance more widely in winter barleys

15   will have significant value for barley growers as well as

significant economic and environmental implications such as

reduced use of fungicide inputs with their associated treatment

costs. The provision of nucleic acid diagnostics as provided

herein enables rapid and accurate deployment of new and

20   existing *mlo* alleles into winter barley germplasm.

Plants which include a plant cell according to the

invention are also provided, along with any part or propagule

thereof, seed, selfed or hybrid progeny and descendants. A

25   plant according to the present invention may be one which does

not breed true in one or more properties. Plant varieties may

be excluded, particularly registrable plant varieties according

35

to Plant Breeders' Rights. It is noted that a plant need not

be considered a "plant variety" simply because it contains

stably within its genome a transgene, introduced into a cell of

the plant or an ancestor thereof.

5   In addition to a plant, the present invention provides any

clone of such a plant, seed, selfed or hybrid progeny and

descendants, and any part of any of these, such as cuttings,

seed. The invention provides any plant propagule, that is any

part which may be used in reproduction or propagation, sexual

10   or asexual, including cuttings, seed and so on. Also

encompassed by the invention is a plant which is a sexually or

asexually propagated off-spring, clone or descendant of such a

plant, or any part or propagule of said plant, off-spring,

clone or descendant.

15   A further aspect of the present invention provides a

method of making a plant cell involving introduction of the

sequence (e.g. as part of a suitable vector) into a plant cell

and causing or allowing recombination between the vector and

the plant cell genome to introduce the sequence of nucleotides

20   into the genome.

Following transformation of a plant cell a plant may be

regenerated.

The invention further provides a method of modulating *Mlo*

25   expression in a plant, which may modulate a defence response in

the plant, comprising expression of a heterologous *Mlo* gene

sequence (or mutant, allele, variant or homologue thereof, as

36

discussed) within cells of the plant. As discussed further
herein, modulation or alteration of the level of constitutive
defence response in a plant may be by way of suppression,
repression or reduction (in the manner of wild-type *Mlo*) or
5   promotion, stimulation, activation, increase, enhancement or
augmentation (in the manner of mutant *mlo*). Activation or
enhancement of the defence response may confer or increase
pathogen resistance of the plant, especially resistance to
powdery mildew and/or rust (such as yellow rust).

10   The term "heterologous" may be used to indicate that the
gene/sequence of nucleotides in question have been introduced
into said cells of the plant or an ancestor thereof, using
genetic engineering, ie by human intervention. A transgenic
plant cell, i.e. transgenic for the nucleic acid in question,
15   may be provided. The transgene may be on an extra-genomic
vector or incorporated, preferably stably, into the genome. A
heterologous gene may replace an endogenous equivalent gene, ie
one which normally performs the same or a similar function, or
the inserted sequence may be additional to the endogenous gene
20   or other sequence. An advantage of introduction of a
heterologous gene is the ability to place expression of a
sequence under the control of a promoter of choice, in order to
be able to influence expression according to preference, such
as under particular developmental, spatial or temporal control,
25   or under control of an inducible promoter. Furthermore,
mutants, variants and derivatives of the wild-type gene, e.g.
with higher or lower activity than wild-type, may be used in

37

place of the endogenous gene. Nucleic acid heterologous, or

exogenous or foreign, to a plant cell may be non-naturally

occuring in cells of that type, variety or species. Thus,

nucleic acid may include a coding sequence of or derived from a

5   particular type of plant cell or species or variety of plant,

placed within the context of a plant cell of a different type

or species or variety of plant. A further possibility is for a

nucleic acid sequence to be placed within a cell in which it or

a homologue is found naturally, but wherein the nucleic acid

10   sequence is linked and/or adjacent to nucleic acid which does

not occur naturally within the cell, or cells of that type or

species or variety of plant, such as operably linked to one or

more regulatory sequences, such as a promoter sequence, for

control of expression. A sequence within a plant or other host

15   cell may be identifiably heterologous, exogenous or foreign.

Down-regulation of wild-type *Mlo* gene function leads to

stimulation of a constitutive defence response. This may be

achieved in a number of different ways, as illustrated below.

The nucleic acid according to the invention may be placed

20   under the control of an inducible gene promoter thus placing

expression under the control of the user.

In a further aspect the present invention provides a gene

construct comprising an inducible promoter operatively linked

to a nucleotide sequence provided by the present invention. As

25   discussed, this enables control of expression of the gene. The

invention also provides plants transformed with said gene

construct and methods comprising introduction of such a

38·

construct into a plant cell and/or induction of expression of a
construct within a plant cell, e.g. by application of a
suitable stimulus, such as an effective exogenous inducer or
endogenous signal.

5   The term "inducible" as applied to a promoter is well
understood by those skilled in the art. In essence, expression
under the control of an inducible promoter is "switched on" or
increased in response to an applied stimulus (which may be
generated within a cell or provided exogenously). The nature of
10   the stimulus varies between promoters. Some inducible promoters
cause little or undetectable levels of expression (or no
expression) in the absence of the appropriate stimulus. Other
inducible promoters cause detectable constitutive expression in
the absence of the stimulus. Whatever the level of expression
15   is in the absence of the stimulus, expression from any
inducible promoter is increased in the presence of the correct
stimulus. The preferable situation is where the level of
expression increases upon application of the relevant stimulus
by an amount effective to alter a phenotypic characteristic.
20   Thus an inducible (or "switchable") promoter may be used which
causes a basic level of expression in the absence of the
stimulus which level is too low to bring about a desired
phenotype (and may in fact be zero). Upon application of the
stimulus, expression is increased (or switched on) to a level
25   which brings about the desired phenotype.

Suitable promoters include the Cauliflower Mosaic Virus
35S (CaMV 35S) gene promoter that is expressed at a high level

39

in virtually all plant tissues (Benfey et al, (1990a) EMBO J 9:
1677-1684); the cauliflower meri 5 promoter that is expressed
in the vegetative apical meristem as well as several well
localised positions in the plant body, eg inner phloem, flower

5   primordia, branching points in root and shoot (Medford, J.I.
(1992) *Plant Cell* 4, 1029-1039; Medford *et al*, (1991) *Plant*
Cell 3, 359-370) and the *Arabidopsis thaliana* LEAFY promoter
that is expressed very early in flower development (Weigel *et*
al, (1992) *Cell* 69, 843-859).

10   An aspect of the present invention is the use of nucleic
acid according to the invention in the production of a
transgenic plant.

When introducing a chosen gene construct into a cell,
certain considerations must be taken into account, well known

15   to those skilled in the art. The nucleic acid to be inserted
should be assembled within a construct which contains effective
regulatory elements which will drive transcription. There must
be available a method of transporting the construct into the
cell. Once the construct is within the cell membrane,

20   integration into the endogenous chromosomal material either
will or will not occur. Finally, as far as plants are concerned
the target cell type must be such that cells can be regenerated
into whole plants.

Plants transformed with the DNA segment containing the

25   sequence may be produced by standard techniques which are
already known for the genetic manipulation of plants. DNA can
be transformed into plant cells using any suitable technology,

40

such as a disarmed Ti-plasmid vector carried by Agrobacterium
exploiting its natural gene transfer ability (EP-A-270355, EP-
A-0116718, NAR 12(22) 8711 - 87215 1984), particle or
microprojectile bombardment (US 5100792, EP-A-444882, EP-A-

5   434616) microinjection (WO 92/09696, WO 94/00583, EP 331083, EP
175966, Green *et al.* (1987) *Plant Tissue and Cell Culture,*
Academic Press), electroporation (EP 290395, WO 8706614) other
forms of direct DNA uptake (DE 4005152, WO 9012096, US
4684611), liposome mediated DNA uptake (e.g. Freeman *et al.*

10   *Plant Cell Physiol.* 29: 1353 (1984)), or the vortexing method
(e.g. Kindle, *PNAS U.S.A.* 87: 1228 (1990d) Physical methods for
the transformation of plant cells are reviewed in Oard, 1991,
Biotech. Adv. 9: 1-11.

    Agrobacterium transformation is widely used by those

15   skilled in the art to transform dicotyledonous species.
Recently, there has been substantial progress towards the
routine production of stable, fertile transgenic plants in
almost all economically relevant monocot plants (Toriyama, et
al. (1988) *Bio/Technology* 6, 1072-1074; Zhang, *et al.* (1988)

20   *Plant Cell Rep.* 7, 379-384; Zhang, *et al.* (1988) *Theor Appl
Genet* 76, 835-840; Shimamoto, *et al.* (1989) *Nature* 338, 274-
276; Datta, *et al.* (1990) *Bio/Technology* 8, 736-740; Christou,
et al. (1991) *Bio/Technology* 9, 957-962; Peng, *et al.* (1991)
International Rice Research Institute, Manila, Philippines 563-

25   574; Cao, *et al.* (1992) *Plant Cell Rep.* 11, 585-591; Li, *et al.*
(1993) *Plant Cell Rep.* 12, 250-255; Rathore, *et al.* (1993)
Plant Molecular Biology 21, 871-884; Fromm, *et al.* (1990)

41

Bio/Technology 8, 833-839; Gordon-Kamm, *et al.* (1990) *Plant Cell* 2, 603-618; D'Halluin, *et al.* (1992) *Plant Cell* 4, 1495-1505; Walters, *et al.* (1992) *Plant Molecular Biology* 18, 189-200; Koziel, *et al.* (1993) *Biotechnology* 11, 194-200; Vasil, I. K. (1994) *Plant Molecular Biology* 25, 925-937; Weeks, *et al.* (1993) *Plant Physiology* 102, 1077-1084; Somers, *et al.* (1992) *Bio/Technology* 10, 1589-1594; WO92/14828). In particular, *Agrobacterium* mediated transformation is now emerging also as an highly efficient alternative transformation method in monocots (Hiei *et al.* (1994) *The Plant Journal* 6, 271-282).

The generation of fertile transgenic plants has been achieved in the cereals rice, maize, wheat, oat, and barley (reviewed in Shimamoto, K. (1994) *Current Opinion in Biotechnology* 5, 158-162.; Vasil, *et al.* (1992) *Bio/Technology* 10, 667-674; Vain et al., 1995, *Biotechnology Advances* 13 (4): 653-671; Vasil, 1996, *Nature Biotechnology* 14 page 702).

Microprojectile bombardment, electroporation and direct DNA uptake are preferred where Agrobacterium is inefficient or ineffective. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, eg bombardment with Agrobacterium coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with Agrobacterium (EP-A-486233).

Following transformation, a plant may be regenerated, e.g. from single cells, callus tissue or leaf discs, as is standard in the art. Almost any plant can be entirely regenerated from

42

cells, tissues and organs of the plant. Available techniques
are reviewd in Vasil et al., *Cell Culture and Somatic Cel
Genetics of Plants, Vol I, II and III, Laboratory Procedures
and Their Applications*, Academic Press, 1984, and Weissbach and

5   Weissbach, *Methods for Plant Molecular Biology*, Academic Press,
1989.

The particular choice of a transformation technology will
be determined by its efficiency to transform certain plant
species as well as the experience and preference of the person

10 practising the invention with a particular methodology of
choice. It will be apparent to the skilled person that the
particular choice of a transformation system to introduce
nucleic acid into plant cells is not essential to or a
limitation of the invention, nor is the choice of technique for

15 plant regeneration.

In the present invention, expression may be achieved by
introduction of the nucleotide sequence in a sense orientation.
Thus, the present invention provides a method of modulation of
a defence response in a plant, the method comprising causing or

20 allowing expression of nucleic acid according to the invention
within cells of the plant. Generally, it will be desirable to
stimulate the defence response, and this may be achieved by
disrupting *Mlo* gene function.

Down-regulation of expression of a target gene may be

25 achieved using anti-sense technology or "sense regulation"
("co-suppression").

In using anti-sense genes or partial gene sequences to

down-regulate gene expression, a nucleotide sequence is placed

under the control of a promoter in a "reverse orientation" such

that transcription yields RNA which is complementary to normal

mRNA transcribed from the "sense" strand of the target gene.

5   See, for example, Rothstein et al, 1987; Smith *et al,* (1988)

Nature 334, 724-726; Zhang *et al,* (1992) *The Plant Cell* 4, 1575-

1588, English *et al.,* (1996) *The Plant Cell* 8, 179-188.

Antisense technology is also reviewed in Bourque, (1995), *Plant*

Science 105, 125-149, and Flavell, (1994) *PNAS USA* 91, 3490-

10   3496.

An alternative is to use a copy of all or part of the

target gene inserted in sense, that is the same, orientation as

the target gene, to achieve reduction in expression of the

target gene by co-suppression. See, for example, van der Krol

15   *et al.,* (1990) *The Plant Cell* 2, 291-299; Napoli *et al.,* (1990)

The Plant Cell 2, 279-289; Zhang *et al.,* (1992) *The Plant Cell*

4, 1575-1588, and US-A-5,231,020.

The complete sequence corresponding to the coding sequence

(in reverse orientation for anti-sense) need not be used. For

20   example fragments of sufficient length may be used. It is a

routine matter for the person skilled in the art to screen

fragments of various sizes and from various parts of the coding

sequence to optimise the level of anti-sense inhibition. It

may be advantageous to include the initiating methionine ATG

25   codon, and perhaps one or more nucleotides upstream of the

initiating codon. A further possibility is to target a

conserved sequence of a gene, e.g. a sequence that is

44

characteristic of one or more genes, such as a regulatory

sequence. Antisense constructs may involve 3'end or 5'end

sequences of *Mlo* or homologues. In cases where several *Mlo*

homologues exist in a plant species, the involvement of 5'- and

5   3'-end untranslated sequences in the construct will enhance

specificity of silencing.

The sequence employed may be about 500 nucleotides or

less, possibly about 400 nucleotides, about 300 nucleotides,

about 200 nucleotides, or about 100 nucleotides. It may be

10   possible to use oligonucleotides of much shorter lengths, 14-23

nucleotides, although longer fragments, and generally even

longer than about 500 nucleotides are preferable where

possible, such as longer than about 600 nucleotides, than about

700 nucleotides, than about 800 nucleotides, than about 1000

15   nucleotides, than about 1200 nucleotides, than about 1400

nucleotides, or more.

It may be preferable that there is complete sequence

identity in the sequence used for down-regulation of expression

of a target sequence, and the target sequence, though total

20   complementarity or similarity of sequence is not essential.

One or more nucleotides may differ in the sequence used from

the target gene. Thus, a sequence employed in a down-

regulation of gene expression in accordance with the present

invention may be a wild-type sequence (e.g. gene) selected from

25   those available, or a mutant, derivative, variant or allele, by

way of insertion, addition, deletion or substitution of one or

more nucleotides, of such a sequence. The sequence need not

45

include an open reading frame or specify an RNA that would be translatable. It may be preferred for there to be sufficient homology for the respective anti-sense and sense RNA molecules to hybridise. There may be down regulation of gene expression even where there is about 5%, 10%, 15% or 20% or more mismatch between the sequence used and the target gene.

Generally, the transcribed nucleic acid may represent a fragment of an Mlo gene, such as including a nucleotide sequence shown in Figure 2, or the complement thereof, or may be a mutant, derivative, variant or allele thereof, in similar terms as discussed above in relation to alterations being made to a coding sequence and the homology of the altered sequence. The homology may be sufficient for the transcribed anti-sense RNA to hybridise with nucleic acid within cells of the plant, though irrespective of whether hybridisation takes place the desired effect is down-regulation of gene expression.

Anti-sense regulation may itself be regulated by employing an inducible promoter in an appropriate construct.

Constructs may be expressed using the natural promoter, by a constitutively expressed promotor such as the CaMV 35S promotor, by a tissue-specific or cell-type specific promoter, or by a promoter that can be activated by an external signal or agent. The CaMV 35S promoter but also the rice *actin1* and maize ubiquitin promoters have been shown to give high levels of reporter gene expression in rice (Fujimoto *et al.*, (1993) *Bio/Technology* 11, 1151-1155; Zhang, *et al.*, (1991) *Plant Cell* 3, 1155-1165; Cornejo *et al.*, (1993) *Plant Molecular Biology*

46

23, 567-581).

For use in anti-sense regulation, nucleic acid including a
nucleotide sequence complementary to a coding sequence of a *Mlo*
gene (i.e. including homologues), or a fragment of a said
5   coding sequence suitable for use in anti-sense regulation of
expression, is provided. This may be DNA and under control of
an appropriate regulatory sequence for anti-sense transcription
in cells of interest.

Thus, the present invention also provides a method of
10   conferring pathogen resistance on a plant, the method including
causing or allowing anti-sense transcription from heterologous
nucleic acid according to the invention within cells of the
plant.

The present invention further provides the use of the
15   nucleotide sequence of Figure 2 or a fragment, mutant,
derivative, allele, variant or homologue thereof, such as any
sequence shown or identified herein, for down-regulation of
gene expression, particularly down-regulation of expression of
an Mlo gene or homologue thereof, preferably in order to confer
20   pathogen resistance on a plant.

When additional copies of the target gene are inserted in
sense, that is the same, orientation as the target gene, a
range of phenotypes is produced which includes individuals
where over-expression occurs and some where under-expression of
25   protein from the target gene occurs. When the inserted gene is
only part of the endogenous gene the number of under-expressing
individuals in the transgenic population increases. The

mechanism by which sense regulation occurs, particularly down-regulation, is not well-understood. However, this technique is well-reported in scientific and patent literature and is used routinely for gene control. See, for example, van

5   der Krol et al., (1990) *The Plant Cell 2*, 291-229; Napoli et al., (1990) The Plant Cell 2, 279-289;   Zhang et al, 1992 *The Plant Cell 4*, 1575-1588.

Again, fragments, mutants and so on may be used in similar terms as described above for use in anti-sense regulation.

10   Thus, the present invention also provides a method of conferring pathogen resistance on a plant, the method including causing or allowing expression from nucleic acid according to the invention within cells of the plant. This may be used to suppress Mlo activity. Here the activity of the product is

15   preferably suppressed as a result of under-expression within the plant cells.

As noted, *Mlo* down-regulation may promote activation of a defence response, which may in turn confer or augment pathogen resistance of the plant, especially resistance to powdery

20   mildew and/or rust (e.g. yellow rust).

Thus, the present invention also provides a method of modulating *Mlo* function in a plant, the method comprising causing or allowing expression from nucleic acid according to the invention within cells of the plant to suppress endogenous

25   *Mlo* expression.

Modified versions of *Mlo* may be used to down-regulate endogenous *Mlo* function. For example mutants, variants,

48

derivatives etc., may be employed. For instance, expression of

a *mlo* mutant sequence at a high level may out-compete activity

of endogenous *Mlo*.

Reduction of *Mlo* wild type activity may be achieved by

5 using ribozymes, such as replication ribozymes, e.g. of the

hammerhead class (Haseloff and Gerlach, 1988, *Nature* 334: 585-

591; Feyter *et al. Mol.*, 1996, *Gen. Genet.* 250: 329-338).

Another way to reduce *Mlo* function in a plant employs

transposon mutagenesis (reviewed by Osborne *et al.*, (1995)

10 *Current Opinion in Cell Biology* 7, 406-413). Inactivation of

genes has been demonstrated *via* a 'targeted tagging' approach

using either endogenous mobile elements or heterologous cloned

transposons which retain their mobility in alien genomes. *Mlo*

alleles carrying any insertion of known sequence could be

15 identified by using PCR primers with binding specificities both

in the insertion sequence and the *Mlo* homologue. 'Two-element

systems' could be used to stabilize the transposon within

inactivated alleles. In the two-element approach, a T-DNA is

constructed bearing a non-autonomous transposon containing

20 selectable or screenable marker gene inserted into an excision

marker. Plants bearing these T-DNAs are crossed to plants

bearing a second T-DNA expressing transposase function. Hybrids

are double-selected for excision and for the marker within the

transposon yielding F_2 plants with transposed elements. The

25 two-element approach has a particular advantage with respect to

Ac/Ds of maize, as the transposed *Ds* is likely to be unlinked

to the transposase, facilitating outcrossing and stabilization

49

of the *Ds* insertion (Jones *et al.,* (1994) *Science* 266, 789-793;

Osborne *et al.,* (1995) *Current Opinion in Cell Biology* 7, 406-

413).

5   The *mlo*-based powdery mildew resistance is caused by the

inactivation of the *Mlo* wild type allele, resulting in a

recessive resistance phenotype. Substances that inhibit the

activity of the Mlo wild type protein may be used to induce the

resistance phenotype.

10   An important hint that complete inactivation of *Mlo*

expression is not essential and may even be detrimental is

provided by the description of mutagen-induced *mlo* resistance

alleles that are likely to have retained residual wild type

allele activity. These alleles exhibit no detectable

15   spontaneous leaf necrosis which negatively affects

photosynthesis rates and yield (Hentrich, W (1979) *Arch.*

Züchtungsvorsch., Berlin 9, S. 283-291).

The Mlo protein is predicted to be membrane-anchored by

seven transmembrane helices (see e.g. Figure 7). This

20   structure prediction has been reinforced by recent analysis of

Mlo homologues in rice and *Arabidopsis thaliana.* Structure

prediction of the *Arabidopsis thaliana* homologue also suggests

the presence of seven transmembrane helices. A comparison of

the *Mlo* homologues revealed in addition conserved cysteine

25   residues in the putative extracellular loops 1 and 3 and high

probabilities of amphipathic helices in the second

intracellular loop adjacent to the predicted transmembrane

50

helices 3 and 4. These conserved structural motifs in the

family of Mlo proteins are reminiscent of G protein coupled

receptors (GPCR) described extensively in mammalian systems.

GPCRs are known to be activated by ligands and to amplify

5   signals intracellularly via heterotrimeric G proteins. Without

in any way providing a limitation on the nature or scope of any

aspect of the present invention, it is predicted that Mlo

activates an inhibitory G alpha subunit of heterotrimeric G

proteins, thus leading to a downregulation of as yet unknown

10   effector proteins.

       The provision herein of Mlo sequence information enables

the identification of antagonists of function of the Mlo

protein (e.g. GPCR function). Antagonists of Mlo may block

receptor activation by its unknown genuine ligand, mimicking

15   recessive mutations in the Mlo gene. Such Mlo antagonists may

be used as crop protection compounds, for example applied

externally to the plant or crop or, where the compound is

peptidyl in nature, delivered internally via a biological

vector (e.g. recombinant infecting viral particle expressing

20   the antagonistic molecule within target plant cells) or via a

transgenic route (plants or plant cells genetically modified to

express the antagonist molecule, perhaps under control of a

promoter inducible by an externally applied compound (eg GST-II

promoter from maize - Jepson et al Plant Molecular Biology

25   26:1855-1866 (1994)) allowing control over the timing of

expresion of the mlo inactivation phenotype.

       Leaf segments of *Mlo* wild type plants may be tested with a

test substance, e.g. from a random or combinatorial compound

library, for resistance upon challenge with pathogen such as

powdery mildew. The detached leaf segment assay is used as a

standard test system to score for susceptibility/resistance

5   upon inoculation with powdery mildew spores. Leaf segments of

7-day-old seedlings of the genotype *Mlo RorI* may be placed on

agar, for example individual wells of 96-well microtiter plates

containing 50μl agar. Different compounds may be applied to

the agar surface in each well at a concentration of about 1ppm

10   dissolved in DMSO. Around seven days after inoculation of the

detached leaf segments with pathogen, such as spores of a

virulent powdery mildew isolate, compounds which induce

resistance may be recognised by the absence of fungal mycelium

on leaf segments in the microtiter plates.

15   A further selection may be used to discriminate between

compounds that act in the mlo pathway and those that confer

resistance by other mechanisms, or those which exhibit a direct

fungitoxic activity. For this purpose mutants in genes (*Ror*

genes) which may be required for *mlo* resistance (Freialdenhoven

20   et al., (1996), The Plant Cell 8, 5-14) may be used. Mutants

of these genes confer <u>susceptibility</u> to powdery mildew attack

despite the presence of <u>mlo</u> resistance alleles. Plants of the

genotype *Mlo rorl* (wild type *Mlo* protein and defective *Rorl*

gene) may be used, for example, to test compounds which induce

25   resistance on *Mlo Rorl* genotypes but exhibit susceptibility on

the *Mlo rorl* genotype, enabling selection of candidate *Mlo*

antagonists. Testing candidate compounds identified using a

52

leaf segment test may be used to drastically reduce the number of candidate compounds for further *in vitro* tests.

A further selection step of candidate antagonists may involve heterologous expression of the Mlo protein or a

5   fragment thereof (e.g. in a baculovirus insect cell system) and subsequent binding assays with labelled molecules. Specific binding of compounds to cell lines expressing wild type Mlo protein is a good indicator of their antagonistic mode of action. Analsis of the deduced Mlo protein sequence has

10   provided strong evidence that the protein is anchored in the membrane via seven transmembrane helices and may represent a novel member of the so-called serpentine receptor family. The conclusion is supported by the sequence data derived from homologous genes identified in barley, rice and *Arabidopsis*.

15   Seven transmembrane proteins have been shown to be expressed at high level in the Baculovirus/insect cell system (up to 10^7 molecules per cell - Tate and Grisshamer, 1996, *TIBTECH* **14**: 426-430). Since the family of Mlo proteins appears to be restricted to the plant kingdom, this provides a low-background

20   environment for compound tests. Candidate compounds which are labelled, radioactively or non-radioactively, may be tested for specific binding to Sf9 insect cells expressing the Mlo protein after infecion with a recombinant baculovirus construct. Specificity of the binding may be tested further by Sf9

25   expression of mutant *mlo* proteins which carry characterised mutations (e.g. as in Table 1) leading *in vivo* to resistance.

53

Thus, in various further aspects the present invention
relates to assays for substances able to interfere with Mlo
function, i.e. confer a *mlo* mutant phenotype, such substances
themselves and uses thereof.

5   The use of Mlo in identifying and/or obtaining a substance
which inhibits Mlo function is further provided by the present
invention, as is the use of Mlo in identifying and/or obtaining
a substance which induces pathogen resistance in a plant.

10   Agents useful in accordance with the present invention may
be identified by screening techniques which involve determining
whether an agent under test inhibits or disrupts Mlo function
to induce an *mlo* phenotype. Candidate inhibitors are
substances which bind Mlo.

15   It should of course be noted that references to "Mlo" in
relation to assays and screens should be taken to refer to
homologues, such as in other species, including rice and wheat,
not just in barley, also appropriate fragments, variants,
alleles and derivatives thereof. Assessment of whether a test
20   substance is able to bind the Mlo protein does not necessarily
require the use of full-length Mlo protein. A suitable
fragment may be used (or a suitable analogue or variant
thereof).

Suitable fragments of Mlo include those which include
25   residues known to be crucial for Mlo function as identified by
mlo mutant alleles (Table 1). Smaller fragments, and analogues
and variants of this fragment may similarly be employed, e.g.

54

as identified using techniques such as deletion analysis or
alanine scanning.

Furthermore, one class of agents that can be used to
disrupt Mlo activity are peptides fragments of it. Such
peptides tend to be short, and may be about 40 amino acids in
length or less, preferably about 35 amino acids in length or
less, more preferably about 30 amino acids in length, or less,
more preferably about 25 amino acids or less, more preferably
about 20 amino acids or less, more preferably about 15 amino
acids or less, more preferably about 10 amino acids or less, or
9, 8, 7, 6, 5 or less in length. The present invention also
encompasses peptides which are sequence variants or derivatives
of a wild type Mlo sequence, but which retain ability to
interfere with Mlo function, e.g. to induce an *mlo* mutant
phenotype. Where one or more additional amino acids are
included, such amino acids may be from Mlo or may be
heterologous or foreign to Mlo. A peptide may also be included
within a larger fusion protein, particularly where the peptide
is fused to a non-Mlo(i.e. heterologous or foreign) sequence,
such as a polypeptide or protein domain.

Peptides may be generated wholly or partly by chemical
synthesis. The compounds of the present invention can be
readily prepared according to well-established, standard liquid
or, preferably, solid-phase peptide synthesis methods, general
descriptions of which are broadly available (see, for example,
in J.M. Stewart and J.D. Young, Solid Phase Peptide Synthesis,
2nd edition, Pierce Chemical Company, Rockford, Illinois

.1984), in M. Bodanzsky and A. Bodanzsky, The Practice of
Peptide Synthesis, Springer Verlag, New York (1984); and
Applied Biosystems 430A Users Manual, ABI Inc., Foster City,
California), or they may be prepared in solution, by the liquid

5   phase method or by any combination of solid-phase, liquid phase
and solution chemistry, e.g. by first completing the respective
peptide portion and then, if desired and appropriate, after
removal of any protecting groups being present, by introduction
of the residue X by reaction of the respective carbonic or

10   sulfonic acid or a reactive derivative thereof.

     Another convenient way of producing a peptidyl molecule
according to the present invention (peptide or polypeptide) is
to express nucleic acid encoding it, by use of nucleic acid in
an expression system, as discussed elsewhere herein. This

15   allows for peptide agents to be delivered to plants
transgenically, by means of encoding nucleic acid. If coupled
to an inducible promoter for expression under control of the
user, this allows for flexibility in induction of an *mlo*
phenotype and pathogen resistance. This may allow for any

20   side-effects arising from interference with Mlo function to be
moderated.

     In one general aspect the present invention provides an
assay method for a substance able to interact with the relevant

25   region of Mlo, the method including:

     (a)   bringing into contact a Mlo polypeptide or peptide
fragment thereoof, or a variant, derivative or analogue

56·

thereof, and a test compound; and

(b)   determining interaction or binding between said

polypeptide or peptide and the test compound.

A test compound found to interact with the relevant

5   portion of Mlo may be tested for ability to modulate, e.g.

disrupt or interfere with, Mlo function, as discussed already

above.

Another general aspect of the present invention provides

10   an assay method for a substance able to induce an *mlo* mutant

phenotype in a plant, the method including:

(a)   bringing into contact a plant or part thereof (e.g.

leaf or leaf segment) and a test compound; and

(b)   determining Mlo function and/or pathogen resistance

15   and/or stimulation of a defence response in the plant.

Susceptibility or resistance to a pathogen may be

determined by assessing pathogen growth, e.g. for powdery

mildew the presence or absence, or extent, of mycelial growth.

Binding of a test compound to a polypeptide or peptide may

20   be assessed in addition to ability of the test compound to

stimulate a defence response in a plant. Such tests may be run

in parallel or one test may be performed on a substance which

tests positive in another test.

25   Of course, the person skilled in the art will design any

appropriate control experiments with which to compare results

obtained in test assays.

Performance of an assay method according to the present
invention may be followed by isolation and/or manufacture
and/or use of a compound, substance or molecule which tests
positive for ability to modulate *Mlo* function and/or induce
5   pathogen resistance, such as resistance to powdery mildew.

The precise format of an assay of the invention may be
varied by those of skill in the art using routine skill and
knowledge. For example, interaction between substances may be
10   studied *in vitro* by labelling one with a detectable label and
bringing it into contact with the other which has been
immobilised on a solid support. Suitable detectable labels,
especially for peptidyl substances include ^{35}S-methionine which
may be incorporated into recombinantly produced peptides and
15   polypeptides. Recombinantly produced peptides and polypeptides
may also be expressed as a fusion protein containing an epitope
which can be labelled with an antibody.

An assay according to the present invention may also take
the form of an *in vivo* assay. The *in vivo* assay may be
20   performed in a cell line such as a yeast strain or mammalian
cell line in which the relevant polypeptides or peptides are
expressed from one or more vectors introduced into the cell.

For example, a polypeptide or peptide containing a
fragment of Mlo or a peptidyl analogue or variant thereof as
25   disclosed, may be fused to a DNA binding domain such as that of
the yeast transcription factor GAL 4. The GAL 4 transcription
factor includes two functional domains. These domains are the

58

DNA binding domain (GAL4DBD) and the GAL4 transcriptional

activation domain (GAL4TAD). By fusing such a polypeptide or

peptide to one of those domains and another polypeptide or

peptide to the respective counterpart, a functional GAL 4

5   transcription factor is restored only when two polypeptides or

peptides of interest interact. Thus, interaction of the

polypeptides or peptides may be measured by the use of a

reporter gene probably linked to a GAL 4 DNA binding site which

is capable of activating transcription of said reporter gene.

10   This assay format is described by Fields and Song, 1989, Nature

340; 245-246. This type of assay format can be used in both

mammalian cells and in yeast. Other combinations of DNA

binding domain and transcriptional activation domain are

available in the art and may be preferred, such as the LexA DNA

15   binding domain and the VP60 transcriptional activation domain.

    When looking for peptides or other substances which

interact with Mlo, the Mlo polypeptide or peptide may be

employed as a fusion with (e.g.) the LexA DNA binding domain,

with test polypeptide or peptide (e.g. a random or

20   combinatorial peptide library) as a fusion with (e.g.) VP60.

An increase in reporter gene expression (e.g. in the case of β-

galactosidase a strengthening of the blue colour) results from

the presence of a peptide which interacts with Mlo, which

interaction is required for transcriptional activation of the

25   β-galactosidase gene.

    The amount of test substance or compound which may be

59

added to an assay of the invention will normally be determined
by trial and error depending upon the type of compound used.
Typically, from about 0.001 nM to 1mM or more concentrations of
putative inhibitor compound may be used, for example from 0.01

5   nM to 100μM, e.g. 0.1 to 50 μM, such as about 10 μM. Greater
concentrations may be used when a peptide is the test
substance. Even a molecule which has a weak effect may be a
useful lead compound for further investigation and development.

Compounds which may be used may be natural or synthetic

10   chemical compounds used in drug screening programmes. Extracts
of plants which contain several characterised or
uncharacterised components may also be used. Antibodies
directed to Mlo or a fragment thereof form a further class of
putative inhibitor compounds. Candidate inhibitor antibodies

15   may be characterised and their binding regions determined to
provide single chain antibodies and fragments thereof which are
responsible for disrupting the interaction. Other candidate
inhibitor compounds may be based on modelling the 3-dimensional
structure of a polypeptide or peptide fragment and using

20   rational drug design to provide potential inhibitor compounds
with particular molecular shape, size and charge
characteristics. It is worth noting, however, that
combinatorial library technology provides an efficient way of
testing a potentially vast number of different substances for

25   ability to interact with and/or modulate the activity of a
polypeptide. Such libraries and their use are known in the
art, for all manner of natural products, small molecules and

60

peptides, among others. The use of peptide libraries may be
preferred in certain circumstances.

Following identification of a substance or agent which
5   modulates or affects *Mlo* function, the substance or agent may
be investigated further. Furthermore, it may be manufactured
and/or used in preparation, i.e. manufacture or formulation, of
a composition for inducing pathogen resistance in a plant.
These may be applied to plants, e.g. for inducing pathogen
10   resistance, such as resistance to powedery mildew. A further
aspect of the present invention provides a method of inducing
pathogen resistance in a plant, the method including applying
such a substance to the plant. A peptidyl molecule may be
applied to a plant transgenically, by expression from encoding
15   nucleic acid, as noted.

A polypeptide, peptide or other substance able to modulate
or interfere with *Mlo* function, inducing pathogen resistance in
a plant as disclosed herein, or a nucleic acid molecule
20   encoding a peptidyl such molecule, may be provided in a kit,
e.g. sealed in a suitable container which protects its contents
from the external environment. Such a kit may include
instructions for use.

25   Further aspects and embodiments of the present invention
will be apparent to those skilled in the art. The present
invention will now be exemplified by way of illustration with

61

reference to the following figures:

Figure 1 Positional Cloning of *Mlo*. The *Mlo* locus has

been mapped with increasing precision on the long arm of barley

5   chromosome 4 using morphological, RFLP and AFLP markers. The

upper part of the figure presents the genetic linkage maps of

these markers relative to *Mlo*. All genetic distances are

indicated in centiMorgan (cM) based on multi-point linkage

analysis except for genetic distances between AFLP markers

10   which are calculated by two-point-estimates. The morphological

marker map (Jørgensen, 1977) positions *Mlo* at a distance of

more than 20 cM to hairy leaf sheath (*Hs*) and glossy

sheath/spike (*gsl*). The RFLP marker map is based on the

analysis of 257 F_2 individuals derived from the cross Carlsberg

15   II *Mlo* Grannenlose Zweizeilige *mlo-11*. The previously

published RFLP map (Hinze *et al.*, 1991) of the same cross was

based on only 44 F_2 individuals. The gene was delimited to a

2.7 cM interval bordered by markers bAO11 and bAL88. AFLP

markers were identified and mapped as described in Experimental

20   procedures. Their genetic distance to *Mlo* is based on the

cross Ingrid *Mlo* x BC_7Ingrid *mlo-3*. The crucial result of the

AFLP analysis has been the identification of two markers, Bpm2

and Bpm9, defining an 0.64 cM interval containing the *Mlo* locus

and one marker (Bpm16) cosegregating with *Mlo* on the basis of

25   more than 4,000 meiotic events. Marker Bxm2 which is located

0.1 cM telomeric to *Mlo* was derived from BAC F15 template DNA

(see below). One YAC clone, YAC YHV303-A6, containing the

62

cosegregating marker Bpm16 and two flanking loci (Bpm2 and

Bpm9), is shown in the middle section of the figure. The

position of marker Bpm9 was only roughly estimated within the

YAC clone as indicated by the arrow. The insert of BAC F15

5   represents a 60 kb subfragment of this YAC as indicated in the

lower part of the Figure. After the identification of AFLP

marker Bpm2 in BAC F15, marker Bxm2 was discovered and

positioned 0.1 cM in telomeric orientation of *Mlo*. The

approximate physical position of AFLP markers Bpm2, Bpm16, and

10   Bxm2 (spanning an interval of approximately 30 kb) as well as

the location of some rare occurring restriction sites are

indicated. Dashed lines below the schematic representation of

BAC F15 DNA show the position of the largest established DNA

sequence contigs. The structure of the *Mlo* gene is given

15   schematically in the bottom line of the Figure. Exons are

highlighted by black boxes. Positions of mutational events are

indicated for the eleven tested *mlo* alleles. Mutant alleles

carrying deletions in their nucleotide sequence are marked with

a Δ; the remaining mutant alleles represent single nucleotide

20   substitutions resulting in amino acid exchanges in each case.

Figure 2 shows an *Mlo* coding sequence and encoded amino

acid sequence according to the present invention. The amino

acid sequence predicted from DNA sequences of RT-PCR products

from Ingrid *Mlo* are shown. Nucleotide numbers are given

25   according to translational start site.

Figure 3 Northern Blot Analysis of *Mlo* Transcript

Accumulation. Total RNA (20 μg) and poly(A)$^+$ RNA (5 μg) of

seven-day-old uninfected barley primary leaves of one wild type
(cultivar Ingrid *Mlo*) and two mutant (BC Ingrid *mlo-1*, BC
Ingrid *mlo-3*) cultivars were isolated, separated on a 1.2%
formaldehyde gel and transferred to a nitrocellulose membrane

5   (Hybond). The filter was probed under stringent conditions
(Sambrook *et al.*, 1989) with the radioactivity labelled full
size RT-PCR product derived from Ingrid *Mlo* (Figure 7). A
clear signal is detected only in the lanes containing poly(A)$^+$
RNA. The signal corresponds to a size of approximately 2 kb.

10   Figure 4 Southern Blot Analysis of Intragenic Recombinants
derived from *mlo* heteroallelic crosses. The alleles of two
RFLP markers flanking *Mlo* on opposite sides of either
susceptible F_2 individuals or homozygous susceptible and
homozygous resistant progeny were determined by Southern blot

15   analysis. Plant DNA (10 µg) of the individuals were digested
with *Pst* I (A) or *Hae* III (B) and hybridized with the
radioactively labelled RFLP markers WG114 (upper panel; maps
3.1 cM in centromeric orientation to *Mlo*; see Figure 1) and
ABG366 (lower panel; maps 0.7 cM in telomeric orientation to

20   *Mlo*; see Figure 1) according to standard procedures (Sambrook
et al., 1989).

A   DNA of the parental lines *mlo-8* and *mlo-1* and two
homozygous susceptible (S, *Mlo Mlo*) and two resistant (R, *mlo
mlo*) progenies derived from two susceptible F_2 plants

25   (designated 1 and 2) were tested. The DNAs in lanes S and R
represent selection F_3 individuals from F_3 families obtained by
selfing the susceptible F_2 individuals 1 and 2. Note that

64

susceptible F$_2$ individuals are expected to be heterozygous at
Mlo in this section scheme. Infection phenotypes were scored
seven days after inoculation with the *mlo* avirulent isolate K1.
DNA from a third susceptible individual of this heteroallelic

5   cross (see Table 7) is not included in this Figure.

B   DNA of the parental lines *mlo-5* and *mlo-1* and seven
homozygous susceptible (S, *Mlo Mlo*) and seven resistant (R, *mlo
mlo*) progeny derived from seven susceptible F$_2$ plants
(designated 1 to 7) were tested. The DNAs in lanes S and R

10   represent selected F$_3$ individuals from F$_3$ families obtained by
selfing the susceptible F$_2$ individuals 1 to 7. DNA was
analyzed from two further susceptible individuals of this
heteroallelic cross only in the F$_2$ generation (8* and 9*).

Figure 5 shows an alignment of genomic sequences covering

15   the barley *Mlo* gene and a rice homologue isolated *via*
crosshybridization with a barley gene specific probe. The top
line shows the barley *Mlo* genomic DNA sequence (exon sequences
underlined). The bottom line shows the rice genomic sequence
containing the rice *Mlo* homologue.

20   **Figure 6** shows an alignment of genomic sequences carrying
the barley *Mlo* gene and a barley homologue isolated *via*
crosshybridization with a barley gene specific probe. The top
line shows the barley *Mlo* genomic DNA sequence (exon sequences
underlined). The bottom line shows the genomic sequence

25   containing the barley *Mlo* homologue.

Figure 7 Nucleotide and Deduced Amino Acid Sequence of the
Barley *Mlo* cDNA. The nucleotide and the deduced amino acid

65

sequence are based on the combined data of RT-PCR and RACE

obtained from experiments using RNA of cultivar Ingrid *Mlo*.

The stop codon is marked by an asterisk, the putative

polyadenylation signal is underlined and the detected termini

5   of RACE products are indicated by arrows above the sequence.

Positions of introns as indentified by comparison with

corresponding genomic clones are labelled by triangles below

the nucleic acid sequence. Six predicted transmembrane

spanning helices according to the MEMSAT algorithm (Jones *et*

10   *al.*, 1994) are boxed in grey colour. A putative nuclear

localization signal (K-K-K-V-R) and casein kinase II site (S-I-

F-D) in the carboxy-terminal half of the protein are shown in

bold type.

Figure 8 shows genomic sequence of rice (Oryza sativa)

15   homologue including coding and flanking sequences.

Figure 9 shows genomic sequence of barley (Hordeum

vulgare) homologue including coding and flanking sequences.

Figure 10 shows cDNA sequence of rice homologue.

Figure 11 shows cDNA sequence of barley homologue.

20   Figure 12 shows cDNA sequence of Arabidopsis thaliana

homologue.

Figure 13 shows amino acid sequence of rice homologue.

Figure 14 shows amino acid sequence of barley homologue.

Figure 15 shows amino acid sequence of Arabidopsis

25   homologue.

Figure 16 shows a pretty box of amino acid sequences of

Mlo, barley, rice and Arabidopsis homologues.

66

All documents mentioned in this document are incorporated
by reference.

EXAMPLE 1 - CLONING OF MLO OF BARLEY

5

Targeted search for AFLP markers tightly linked to Mlo

Efforts to increase the DNA marker density around *Mlo* were
coordinated with attempts to construct a local high resolution
genetic map. An alternative possibility would have been to
10   extend the population size of the characterized cross Carlsberg
II *Mlo* x Grannenlose Zweizeilige *mlo-11* (Hinze *et al.*, 1991)
but it was felt to be advantageous to establish a high
resolution map starting out from one of the available
BC *mlo* lines and its recurrent parent line. Importantly, the
15   donor parent of the BC line represents a different genetic
background in comparison to the recurrent parent line. In this
way, searching for linked AFLP markers could be started in
parallel with generating a large mapping population from a
cross between the same genetic lines. In addition, the BC line
20   based cross allowed testing of colinearity of DNA markers in
the vicinity of *Mlo* as determined from the cross
Carlsberg II *Mlo* x Grannenlose Zweizeilige *mlo-11* (Hinze *et*
al., 1991). For the new cross a *mlo-3* backcross (BC) line was
used that had been backcrossed seven times into the genetic
25   background Ingrid (BC$_7$ Ingrid *mlo-3* ; Hinze *et al.*, 1991). The
line was previously characterized to carry a relatively small
introgressed DNA segment on barley chromosome 4. In addition,

the donor parent line Malteria Heda *mlo-3* exhibits in

comparison to DNA from the recurrent parent Ingrid

polymorphisms with most of the identified RFLP loci linked to

Mlo. Thus, by searching polymorphisms only between two DNA

5   templates, from lines Ingrid *Mlo* and BC_7 Ingrid *mlo-3*, we hoped

to increase the density of DNA markers with AFLPs around *Mlo* in

a targeted manner.

The same two lines were crossed to establish a segregating

population for high resolution mapping of DNA markers, formally

10   representing an eigth backcross. F_2 individuals were scored for

mlo resistance after powdery mildew inoculation with isolate Kl

(virulent on Ingrid *Mlo* and avirulent on BC_7 Ingrid *mlo-3*).

Initially, only a small fraction of the F_2 (77 individuals) was

analyzed for recombination events with flanking RFLP markers.

15   Analysis of four identified recombinants (designated 8-32-2, 7-

38-4, 1-34-1, and 1-49-4) indicated colinearity of marker order

in this cross compared to the previously analyzed cross

Carlsberg II *Mlo* x Grannenlose Zweizeilige *mlo-11* (Hinze *et*

al., 1991). Several of the 77 F_2 seedlings which exhibited a

20   susceptible phenotype and heterozygosity for the tested

flanking DNA marker loci (bA011, bAL88/2, and bAP91; Hinze *et*

al., 1991) were grown to maturity to provide further selfed

seed material segregating for *Mlo/mlo-3* in the F_3 generation.

In total, leaf material was harvested for high resolution

25   marker mapping from 2,026 individuals derived from either the

selfed F_2 or F_3 generation.

AFLP marker candidates were identified by testing all

68

possible *Pst* I/*Mse* I primer combinations (1,024) extending into
genomic sequences up to nucleotide positions +2 and +3,
respectively. Similarly, almost 1,900 *Eco* RI/*Mse* I primer
combinations (+3/+3) have been analyzed. Four DNA templates

5   were included in this analysis: Ingrid *Mlo*, BC$_7$ Ingrid *mlo-3* , a
DNA pool of two phenotypically *mlo* resistant F$_2$ individuals,
and a DNA pool of nine phenotypically susceptible F$_2$
individuals. The resistant and susceptible F$_2$ individuals which
were included as DNA pools in the AFLP search had been selected

10   from the above mentioned RFLP analysis of 77 F$_2$ segregants. The
pooled F$_2$ DNA enabled us to control whether candidate
polymorphisms detected between template DNA from the parents
were heritable traits in the F$_2$. All identified AFLP candidate
markers have been re-examined with eight DNA templates: Ingrid

15   *Mlo*, BC$_7$ Ingrid *mlo-3* , DNA pools from individuals of three F$_3$
families which were phenotypically homozygous susceptible
(*MloMlo*) according to K1 inoculation experiments; DNA of three
resistant F$_2$ individuals. A total of 18 *Pst* I/*Mse* I and 20
Eco RI/*Mse* I primers were confirmed based on the selection

20   procedure.

    The number of identified AFLP markers made it useful to
assign them first roughly to marker intervals based on the RFLP
map around *Mlo*. It was hoped that this approach should enable
both evaluation of the distribution of AFLPs among previously

25   identified RFLP intervals close to *Mlo* and selection of a pair
of flanking AFLP markers with which recombinants could be
identified among the 2,026 segregants. For AFLP assignment we

69

used those four recombinants that had been identified with RFLP

markers out of the above mentioned small sample of 77 F_2

segregants from Ingrid Mlo x BC_7 Ingrid $mlo-3$ (two recombinants

in interval bAP91-bAL88, one in Mlo-bAO11, and one in bAO11-

5   ABG366). A total of 18 AFLPs were found to be located within a

genetic distance of approximately 3.5 cM including Mlo.

Construction of a high resolution AFLP map around Mlo

A two-step procedure was used to construct the high

10   resolution AFLP map. First, all 2,026 segregants were screened

for recombination events between two AFLP markers on opposite

sides of Mlo and subsequently only the few identified

recombinants were used to map all the identified AFLPs in the

3.5 cM target interval. AFLP markers Bpml and Bpm9 were

15   chosen, detecting each allelic DNA fragments in Ingrid Mlo and

BC_7 Ingrid $mlo-3$ and located on opposite sites of Mlo to screen

DNA templates of the segregants for recombination events.

Alternatively, the search for recombinants could have been

carried out with the flanking RFLP markers bAO11 and bAL88.

20   However, although the conversion into cleaved amplified

polymorphic sites (CAPS) was successful for both markers,

difficulties to display the alleles of both loci simultaneously

from crudely purified genomic DNA were encountered. A total of

2,026 individuals (F_2 or F_3 segregants) were screened

25   simultaneously with AFLP markers Bpml and Bpm9 and 98

recombinants were identified. AFLP analysis was subsequently

carried out with each of the 98 DNA templates of the

recombinants to identify the alleles of each of the identified

of AFLP loci. The recombinants have been selfed and

inoculation experiments with powdery mildew isolate K1 were

performed using at least 25 individuals of each recombinant

5   family to deduce the alleles of the previous generation at the

Mlo locus. The obtained data enabled the construction of a high

resolution map around *Mlo* based on more than 4,000 meiotic

events and a resolution of at least 0.025 cM derived *via* two-

point estimates. The essential result has been the

10   identification of a DNA marker cosegregating with *Mlo* (Bpm16)

and two flanking markers (Bpm2 and Bpm9) at a distance of 0.25

and 0.4 cM respectively (Figure 1).

Construction of a large insert size barley YAC library,

15   *isolation of Bpm16 containing YACs, and physical delimitation*

of Mlo

      The genetic evidence indicates that *mlo* resistance is due

to loss of function in the *Mlo* wild type allele. Therefore, it

was decided to establish a large insert size YAC library from

20   cultivar Ingrid *Mlo* into vector pYAC4 (Burke *et al.*, 1987;

Hieter, 1990). Megabase DNA suitable for YAC cloning

experiments was prepared in mg amounts from mesophyll

protoplasts of five-day-old seedlings according to a modified

protocol described by Siedler and Graner (1991). The DNA was

25   partially digested with *Eco* RI in the presence of *Eco* RI

methyltransferase to obtain DNA fragments after preparative

pulsed-field gel electrophoresis (PFGE) in the size range of

71

500-600 kb. After ligation with *Eco* RI digested pYAC4, the DNA was transformed into yeast strain AB1380 and colonies carrying recombinant pYAC4 DNA were selected on solidified synthetic complete medium lacking tryptophan and uracil (Sherman *et al.*,

5   1986). Forty randomly selected yeast colonies were tested for the presence of barley DNA using labelled barley genomic DNA in Southern experiments. The size of the YAC inserts was found after PFGE separations to vary between 500 and 800 kb. On average a genetic distance of 0.2 cM was expected to be

10   represented on the individual recombinant YAC clone. A total of ~40,000 clones representing four barley genome equivalents have been generated.

Four YAC clones (designated 303A6, 322G2, 400H11, and 417D1) have been isolated with marker Bpm16 cosegregating with

15   *Mlo*. Their insert size was determined by PFGE to be 650, 710, 650, and 820 kb respectively. AFLP analysis had shown that three of these clones (303A6, 322G2, and 417D1) contain also both flanking marker loci whereas clone 400H11 contains only loci Bpm16 and Bpm2. These findings strongly suggested that the

20   *Mlo* gene had been physically delimited on recombinant YAC clones 303A6, 322G2, and 417D1.

YAC 303A6 was chosen for subcloning experiments into BAC vector pECSBAC4 containing a unique *Eco* RI site (Shizuya *et al.*, 1992; the vector pECSBAC4 is described by Frijters and

25   Michelmore, 1996; submitted). Total yeast DNA of this clone was partially digested with *Eco* RI to obtain DNA fragments with an average size of 50 kb and ligated into *Eco* RI digested and

72

dephosphorylated BAC vector. Bacterial colonies containing

YAC 303A6-derived DNA in pECSBAC4 were identified by replica

colony hybridization experiments. One set of colony containing

membranes was hybridized with labelled yeast AB1380 DNA and the

5   replica set was hybridized with labelled PFGE-purified YAC303A6

DNA. Recombinant BAC clones containing the AFLP locus Bpm16

were subsequently identified using the cloned 108 bp

Pst I/*Mse* I genomic Bpm16 fragment as a probe in colony

hybridization experiments.

10   . One BAC clone, BAC F15, containing an insert of ~ 60 kb

was chosen for further detailed studies. It was found that the

recombinant BAC clone contained in addition the AFLP marker

locus Bpm2, but not Bpm9. At this point the BAC F15 insert DNA

indicated successful physical delimitation in telomeric

15   orientation but it was an open question whether the insert

would contain bordering sequences in centromeric direction.

Instead of constructing a BAC contig between Bpm 16 and Bpm9,

the option to develop new polymorphic markers from BAC F15 was

chosen. An allelic *Xba* I/*Mse* I polymorphism (designated Bxm2)

20   was identified between the parental lines Ingrid *Mlo* and

BC₇ Ingrid *mlo-3*.

An analysis of the 25 recombinant individuals carrying

recombination events within the *Mlo* containing interval Bpm2-

Bpm9 enabled mapping of Bxm2 in centrometric orientation at a

25   distance of 0.1 cM from *Mlo*. Only four out of the 16 available

recombinants in the interval Bpm9-*Mlo* and none of the 9

recombinants in the interval *Mlo*-Bpm2 were found to exhibit a

73

recombination event between Bxm2 and *Mlo*. It was concluded

that *Mlo* had been physically delimited on BAC F15 between

marker loci Bpm2 and Bxm2 (Figure 1).

5　　*Identification of the Mlo gene and mlo mutants*

A random sequencing project was initiated to determine

sequence contigs of the ~60 kb insert of BAC F15 before marker

Bxm2 was identified and shown to delimit the gene in telomeric

orientation. In parallel, a physical map was generated

10　　(Figure 1). The physical map indicated that the flanking

markers Bpm2 and Bxm2 are physically separated by ~30 kb. The

sequence contigs were searched for regions of high coding

probability using the UNIX versions of the STADEN program

package. Only one sequence contig of almost 6 kb, including the

15　　cosegregating marker Bpm16, revealed an extensive region of

high coding probability.

RT-PCR reactions were performed with total leaf RNA

derived from cultivar Ingrid *Mlo* using a series of primers

deduced from regions which indicated high coding probabilities

20　　and obtained in each case a distinct amplification product.

Sequencing of the largest RT-PCR products revealed a single

extensive open reading frame of 1,602 bp (Figure 2). The

deduced putative protein of 533 amino acids has a molecular

weight of 60.4 kDal. The ~1.7 kb RT-PCR product was used as a

25　　hybridization probe and detected a single RNA transcript of

~1.9 kb length. (Figure 3). A comparison of the genomic

sequence and the largest RT-PCR fragment reveals 12 exons and

74

11 introns, each flanked by the characteristic splice site
sequences (Figure 1).

5   Because marker Bpm16 is located at the 3' end of the above
     described gene (exon 11) and cosegregates with the Mlo locus,
     we started a direct PCR sequencing of the various available
     mutagen-induced mlo resistance alleles. We identified in 14
     out of 15 tested mutant alleles nucleotide alterations which
     result either in single amino acid alterations, deletions or
10   frame shifts of the wild type sequence (Table 1). We suspect
     that mutant allele mlo-2 is located within the promoter- or 5'
     untranslated sequences. The region is notoriously difficult to
     be sequenced via direct PCR sequencing from genomic DNA
     templates but experiments using a series of nested primers are
15   likely to solve this problem. In summary, the comparative
     sequencing of genomic DNA from various mutant mlo lines and
     their respective Mlo wild type ciltivars provided strong
     evidence that Mlo has been identified.

20   Intragenic recombinants
         It had been the intention to provide a chain of evidence
     for the molecular isolation of Mlo which did not rely upon
     complementation experiments via transgenic barley plants. We
     had chosen to develop an unusual genetic tool to confirm that
25   the identified gene represented Mlo. It was reasoned that if
     the mutations observed in the above described gene caused
     resistance to the powdery mildew fungus, recombination events

75

between mutant allele sites should restore wild type sequences.
It was predicted that those intragenic recombinants would
exhibit susceptibility upon powdery mildew attack.

A crossing scheme was devised involving *mlo* resistance

5   alleles *mlo-1*, *mlo-5*, and *mlo-8*. The mutant alleles originate
from the genetic backgrounds Haisa (*mlo-1*) and Carlsberg II
(*mlo-5* and *mlo-8*). Intermutant crosses were performed as shown
in Table 2 generating in each case at least 10 F_1 plants. F_2
populations were obtained by self-fertilization. F_2 seedlings

10   were screened for rare susceptible individuals after
inoculation with powdery mildew isolate K1 which is virulent on
each of the parental *Mlo* wild type cultivars. Susceptible F_2
individuals were identified with a frequency of ~6 x 10^{-4}. In
contrast, if comparable numbers of progenies from selfings of

15   each of the *mlo* mutants were tested for resistance to K1, no
susceptible seedling was identified. This finding strongly
indicated that the majority of the susceptible individuals
derived from the intermutant crosses were not due to
spontaneous reversion events of the mutant *mlo* alleles.

20   Inheritance of the susceptible F_2 individuals was tested
after selfing in F_3 families. Each of the F2 individuals
segregated susceptible and resistant F_3 individuals indicating
hetrozygosity for alleles conferring resistance/susceptibility
in the F_2. Homozygous susceptible F_3 progeny were isolated for

25   the majority of susceptible F_2 individuals by selfing of F_3
individuals and subsequent identification of F_4 families in
which only susceptible individuals were detected.

76

A molecular analysis of the susceptible individuals has
been performed using RFLP markers known to be tightly linked
(< 3 cM) on each side of the *Mlo* locus (Figure 4). RFLP marker
WG114 maps in centromeric orientation relative to *Mlo*, marker

5   ABG366 maps in the direction of the telomere. Detected RFLP
alleles are shown for the intermutant crosses *mlo-8* x *mlo-1* (A)
and *mlo-1* x *mlo-5* (B). DNA was analyzed either from susceptible
F_2 individuals (indicated by *) or from homozygous susceptible
(S) and homozygous resistant (R) F_3 progeny obtained from

10   selfed susceptible F_2 individuals.

The homozygous susceptible F_3 progeny from the susceptible
F_2 plant #1 of cross
mlo-8 x *mlo-1* (Figure 4) reveals the WG114 allele derived from
the *mlo-1* parent in centromeric orientation next to *Mlo* and the

15   ABG366 allele from the *mlo-8* parent in telomeric orientation to
Mlo. The homozygous resistant F_3 progeny from F_2 plant #1 of
this cross reveals in contrast only the flanking marker alleles
derived from parent *mlo-1*. The finding strongly suggested that
susceptibility in F_2 plant #1 is caused by a cross-over type of

20   recombination in the preceding meiosis of one chromosome which
results in a restoration of the *Mlo* wild type allele whereas
the second F_2 chromosome of individual 1 contains a
functionally unaltered *mlo-1* allele. The allelotypes of the
RFLP loci of the homozygous susceptible F_3 progeny from

25   susceptible F_2 plant #2 are identical to the one described
above. However, flanking marker alleles from the homozygous
resistant F_3 progeny of this individual are in both cases

77

derived from the *mlo-8* parent. It is concluded that again a cross-over type of recombination restored one *Mlo* wild type allele in the susceptible F_2 individual #2.

Nine susceptible F_2 individuals were recovered from the
5   cross *mlo-1* x *mlo-5* (Figure 4). For susceptible F_2 individuals #1 to #7 both homozygous susceptible and homozygous resistent F_3 progeny were analyzed at the DNA level. Note that only DNA from the heterozygous susceptible F_2 individuals was analyzed in the case of individuals #8 and #9 (marked by a *). The
10 following allele patterns with respect to the flanking RFLP loci were observed: (*i*) homozygous resistant F_3 progeny showed on both sides of *Mlo* either only the allelotypes of loci WG114 and ABG366 derived from the *mlo-1* parent (individuals #1, #3, #6, #7) or only the allelotypes derived from the *mlo-5* parent
15 (individuals #2, #4, #5). (*ii*) Homozygous susceptible F_3 progeny showed in contrast either only the allelotypes of both loci derived from the *mlo-5* parent (no. #3, #5, #6) or they showed different allelotypes on both sides of *Mlo* (individuals #1, #2, #4, #7). (*iii*) The homozygous susceptible F_3 progeny
20 with different allelotypes on both sides always contain in centromeric orientation the *mlo-1* derived WG114 allele and in telomeric orientation the *mlo-5* derived ABG366 allele. (*iv*) The heterozygous susceptible F_2 individual #8 reveals on either side next to *Mlo* only the alleles derived from parent *mlo-5*.
25 The heterozygous susceptible individual #9 reveals in centromeric orientation alleles derived from both parents *mlo-1* and *mlo-5* whereas only the *mlo-5* derived allele is detected in

78

telomeric orientation. A comprehensive interpretation of the
data suggests that susceptibility in F_2 individuals no. #1, #2,
#4, #7, and #9 is caused by a cross-over type of recombination
restoring the Mlo wild type allele. Non cross-over types of

5   recombination may have restored the Mlo wild type allele in
individuals no. #3, #5, #6, and #8.

A compilation of the detected flanking RFLP alleles of all
isolated susceptible F_2 individuals or homozygous F_3 progeny is
shown in Table 3. Note that individual #3 of the cross mlo-8 x

10   mlo-1 is not shown in Figure 4. The compilation reveals that
(i) cross-over types of recombination (CO) and non cross-over
types of recombination (NCO) are found with a ratio of 7 : 5,
(ii) cross-over types of recombination are resolved
unidirectional, and (iii) NCO recombinants were not observed

15   with parental mlo-1-linked RFLP alleles.

The CO type intragenic recombinants isolated from
heteroallelic mlo crosses were used to test whether wild type
sequences of the Mlo candidate gene had been restored. For the
three relevant alleles mlo-1, mlo-5, mlo-8 alleles candidate

20   mutation sites have been identified (Table 1 and 4). Direct
PCR sequencing of genomic DNA of susceptible intragenic
recombinants derived from both heteroallelic crosses mlo-1 x
mlo-8 and mlo-1 x mlo-5 revealed restoration of wild type
sequences (Table 4). This observation strongly suggests that

25   the intragenic cross over event occurred between nucleotide -1
and +483 in the former and +3 and +483 in the latter cross
(according to translational start site). Thus, the molecular

analysis of seven intragenic recombinants from two

heteroallelic crosses provides final proof that the above

described candidate gene represents *Mlo*.

5   *EXAMPLE 2 - HOMOLOGUES OF THE IDENTIFIED MLO GENE*

The available expressed sequence tag (EST) databases of

Oryzae sativa (rice) and *Arabidopsis thaliana* were searched for

homologous protein sequences. Five Arabidopsis cDNA clones were

10   identified whose deduced amino acid sequences show substantial

similarity to the Mlo protein. Remarkable is cDNA clone

205N12T7 which reveals a chance probability of $1.2\ e^{-45}$. In

addition, at least one significant homologue was found in rice

(OSR16381A).

15   A rice BAC library (Wang *et al.*, 1995) has also been

screened with a labelled barley genomic fragment containing

Mlo. A BAC clone containing an insert of ~23 kb was isolated.

Subsequent subcloning enabled isolation of a 2.5 kb *Pst* I

genomic rice fragment showing strong cross-hybridization with

20   the barley *Mlo* gene probe. DNA sequencing of this fragment

revealed remarkable DNA sequence similarities within exon

sequences of the barley *Mlo* gene (Figure 5).

Finally, a 13 kb λ genomic barley clone derived from

cultivar Igri (Stratagene) was isolated with a labelled barley

25   genomic fragment containing *Mlo*. The nucleotide sequence

derived from a subcloned 2.6 kb Sac I fragment reveals again

extensive sequence similarities to the *Mlo* gene (Fig. 6). The

80

location of the barley *Mlo* homologue within the genome is not
within BAC F15 DNA.

In summary, there is conclusive evidence for *Mlo*
homologues both in a monocotyledonous and a dicotyledonous

5   plant species.

Discussion

Any speculation as to mode of action of *Mlo* and *mlo*
nucleic acid and polypeptides should provide no limitation on

10   the nature or scope of any aspect or embodiment of the present
invention.

In plants, resistance to pathogens is frequently
determined by dominant resistance genes, whose products are
assumed to recognize pathogen-derived avirulence gene products.

15   This mode of pathogen defence follows Flor's gene-for-gene
hypothesis (Flor, 1971). Recently, several 'gene-for-gene' type
resistance genes have been molecularly isolated (Martin *et al.*,
1993; Bent *et al.*, 1994; Jones *et al.*, 1994; Mindrinos *et al.*,
1994; Whitham *et al.*, 1994; Grant *et al.*, 1995; Lawrence et

20   *al.*, 1995; Song *et al.*, 1995). The surprising finding is that
the deduced proteins share remarkable similar structural
domains although they trigger resistance reactions to pathogens
such as viruses, fungi, and bacteria (Dangl, 1995; Staskawicz
et al., 1995). The isolated genes code for proteins that either

25   contain a leucine-rich region (LRR), with or without an
attached nucleotide binding site (NBS), indicative of ligand-
binding and protein-protein interaction or encode a simple

81

serine/threonine kinase. A structural combination of LRR and

the kinase domain has been reported in the deduced protein from

the rice *Xa21* resistance gene (Song *et al.*, 1995). The

structural similarity of resistance genes in 'gene-for-gene'

5   defence makes the existence of a common underlying resistance

mechanisms likely.

Resistance mediated by recessive resistance alleles of the

Mlo gene differs in various aspects from 'gene-for-gene'

resistance (see introductory comments above). The molecular

10   isolation of the *Mlo* gene and the sequencing of various

mutation-induced *mlo* alleles described here, confirms previous

interpretations from combined mutational and Mendelian genetic

studies (Hentrich, 1979; Jørgensen, 1983). It is concluded that

defective alleles of the *Mlo* locus mediate broad spectrum

15   resistance to pathogens such as the powdery mildew pathogen.

This is inconsistent with the involvement of a specific

recognition event of a pathogen-derived product as has been

proposed for race-specific resistance genes.

Pleiotropic effects of *mlo* alleles have provided some

20   clues towards the development of a molecular concept of the

observed broad spectrum resistance response.

Firstly, aseptically grown *mlo* plants exhibit at a high

frequency a spontaneous formation of cell wall appositions

(CWAs) in leaf epidermal cells (Wolter *et al.*, 1993). Those

25   CWAs are usually formed in response to attempted pathogen

penetration directly beneath the fungal apressorium. CWAs are

believed to form a physical barrier against pathogen ingress

82

and have been implicated repeatedly in *mlo* mediated resistance
(Bayles, 1990).

Secondly, at a later stage, the plants develop
macroscopically detectable leaf necrotic flecks. The

5   spontaneous leaf necrosis response has been extensively studied
with a unique collection of 95 chemically-induced *mlo* alleles
(Hentrich, 1979). The alleles were classified as either showing
a gradually different infection phenotype upon infection of a
mixture of nine powdery mildew isolates. Those *mlo* alleles

10   which give rise to an intermediate infection phenotype (i.e.
development of a considerable number of sporulating fungal
colonies upon inoculation) showed no detectable spontaneous
leaf necrosis whereas the category of the most effective
resistance alleles exhibits pronounced necrosis in the absence

15   of the pathogen. Thus, there is solid evidence that the former
category of *mlo* alleles retain residual wild type allele
activity and those alleles appear to exhibit no detectable
spontaneous leaf necrosis.

Thirdly, a constitutive expression of defence-related

20   genes has been observed in *mlo* seedlings grown under mildew-
free conditions - in primary leaves when 10-11 days old; this
includes genes of the *PR-1* family, chitinases and peroxidases.

We have shown that *mlo* in barley confers increased
resistance to different types of yellow rust (*Puccinia*

25   *struciformis*) when a one to one mixture of talcum powder and
spores were aviblown onto leaves of *mlo* barley plants after
onset of constitutive expression of defence related genes (10-

83

11 day old *mlo* seedlings).

Thus, it appears that multiple defence-associated responses are constitutively expressed in *mlo* plants.

The temporal relationship of these events is interesting: the onset of constitutive defence-related transcript accumulation is detected in 11 day-old seedlings and precedes CWA formation which is followed by the appearance of macroscopically visible leaf necrosis. Importantly, however, *mlo* resistance can be experimentally tested as early as in five day-old seedlings and is fully functional at this time. We conclude that the Mlo protein has a negative regulatory function in plant defence and that plants with a defective protein are 'primed' for the onset of defence responses.

The deduced amino acid sequence of *Mlo* reveals no significant homologies to any of the described plant resistance genes so far, supporting the idea of a distinct molecular resistance mechanism. The *Mlo* gene shows also no striking similarities to any characterized plant or mammalian gene sequence in the various data bases. However, highly significant homologous sequences have been identified in the EST and genomic databases both from rice and *Arabidopsis thaliana* (Table 5 and Figure 5). This strongly suggests that the Mlo protein represents a member of a novel protein family. A putative nuclear localization motif (NLS) is found within exon 12 providing indication of nuclear localization of the protein (KEKKKVR; Nigg *et al.*, 1991). The significance of this motif is supported by a casein kinase II motif located 14 amino acids

84

into direction of the NH$_2$-terminus (SIFD; Rihs et al., 1991).
Functional tests may examine the putative subcellular
localization of the Mlo protein.

Mutations have been described also in other plant species
5   in which defence responses to pathogens appear to be
constitutively expressed (Walbot et al., 1983; Pryor, 1987;
Jones, 1994). It has been suggested that this class of mutants,
termed lesion mimics (Les) or necrotic mutants (nec), affect
the control of plant defence responses. Recessively inherited
10   lesion mimic mutants have been systematically analysed in
Arabidopsis thaliana (Greenberg and Ausubel, 1993; Dietrich et
al., 1994; Greenberg et al., 1994; Weymann et al. 1995). The
affected genes have been designated acd (accelerated cell
death; acd1 and acd2) or lsd (lesions simulating disease
15   resistance response; lsd1 to lsd7).

Each of the mutants exhibits, in the absence of pathogens,
HR characteristics such as plant cell wall modifications and
the accumulation of defence-related gene transcripts. Leaves of
the acd2 mutant have been shown to accumulate high levels of
20   salicylic acid and of the Arabidopsis phytoalexin, camelexin
(Tsuji et al., 1992). Importantly, acd and lsd mutants exhibit
elevated resistance to a bacterial (P. syringae) and fungal (P.
parasitica) pathogen. The lsd1 mutant is exceptional in that it
confers heightened pathogen resistance at a prelesion state, in
25   contrast to the other defective loci which exhibit elevated
pathogen resistance only in the lesion-positive state. In this
respect, lsd1 resembles the mlo mutants in barley. Another

striking feature of *lsd1* is the indeterminate spread of lesions in contrast to the other mutants where lesion growth is determinate.

5   *EXPERIMENTAL PROCEDURES*

Plant Material

A compilation of the *mlo* mutants and their mother varieties analyzed in this study has been described by

10 Jørgensen (1992) [*mlo-1*, *mlo-3*, *mlo-4*, *mlo-5*, *mlo-7*, *mlo-8*, *mlo-9*, *mlo-10*, *mlo-11*] and by Habekuss and Hentrich (1988) [mutants in cultivar Plena 2018 (*mlo-13*), 2034 (*mlo-17*), 2118]. Since mutant 2118 has not been assigned to an allele number so far, we designate the allele here as *mlo-26*, according to

15 current numbering in the GrainGene database (gopher://greengenes.cit.cornell.edu:70/77/.graingenes.ndx/ index?mlo).

The high resolution map is based on a cross between Ingrid *Mlo* x BC$_7$ Ingrid *mlo-3*. F$_1$ plants were selfed generating a

20 segregating F$_2$ population of approximately 600 plants. Phenotypically susceptible F$_2$ plants which showed heterozygosity for RFLP markers on opposite sites of *Mlo* were selfed and generated further segregants in the F$_3$ generation for high resolution mapping.

25

Powdery Mildew Infection Tests

The fungal isolate K1 (Hinze et al., 1991) is virulent on

86

all cultivars used in this study carrying the *Mlo* allele and
avirulent on all tested *mlo* genotypes. Plant growth and
inoculation with *Erysiphe graminis* f sp *hordei* were carried out
as described previously (Freialdenhoven et al., 1996). The
5   genotype at *Mlo* of recombinants used for the high resolution
map were determined after selfing and subsequent inoculation
experiments in F_3 or F_4 families comprising at least 24
individuals.

10   *AFLP Analysis*

Genomic DNA for AFLP analysis was isolated according to
Stewart and Via (1993). AFLP analysis was carried out with
minor modifications as described by Vos et al. (1995). For
screening of AFLP markers linked to *Mlo* we used the enzyme
15   combinations Pst I/Mse I with amplification primers carrying +2
and +3 selective bases respectively in genomic sequences of
amplified fragments. For Eco RI/Mse I amplification primers we
used +3 and +3 selective bases respectively. A set of four DNA
templates has been used: from the susceptible parent cultivar
20   Ingrid *Mlo*, the resistant parent BC_7Ingrid *mlo*-3, a pool of two
resistant F_2 individuals (*mlo*-3 *mlo*-3) and a pool of nine
susceptible F_2 individuals (*Mlo Mlo*) derived from the cross
Ingrid *Mlo* x BC_7 Ingrid *mlo*-3. Amplified genomic fragments
representing AFLP markers Bpm2, Bpm9, and Bpm16 (Figure 1) were
25   cloned and sequenced as follows: gel pieces (fixed by vacuum
drying to Whatman 3MM paper) containing the amplified genomic
fragments were identified via autoradiography and subsequently

87

excised.   100 μl water were added, boiled for 10 min. and after

centrifugation 5 μl of the supernatant were used as a template

for non-radioactive reamplification (30 cycles) with the

selective AFLP primers.   Amplification products were isolated

5   after agarose gel using a DNA isolation kit (Jetsorb, Genomed

Inc., USA).   DNA was reated with Klenow polymerase and T4

polynucleotide kinase and subsequently cloned in the EcoRV site

of pBluescript SK (Stratagene).   Sequencing reactions were

performed using a dye terminator cycle sequencing reaction kit

10   (Perkin Elmer) and resolved either on an ABI 373 or 377

(Applied Biosystems) automated sequencer.

Barley YAC Library and BAC Sublibrary Construction of YAC
YHV303-A6

15   The YAC library of barley cultivar Ingrid was established

using the pYAC4 vector (Burke et al., 1987; Kuhn and Ludwig

1994) and yeast strain AB 1380.   Details of the library

construction and its characterization will be described

elsewhere.   Screening for YAC clones containing marker Bpm16

20   was done by AFLP analysis.   For construction of a BAC

sublibrary of YAC YHV303-A6, total DNA of this yeast clone was

used.   After partial Eco RI digestion and preparative pulsed-

field gel electrophoresis, DNA fragments in the size range of

50 kb were recovered and subcloned in the pECSBAC4 vector.

25   Clones carrying YHV303-A6 derived inserts were identified by a

two-step colony hybridization procedure.   First total labelled

DNA of the non-recombinant yeast strain AB 1380 was used as a

88

probe to eliminate most of the clones carrying insert DNA

derived from the host strain. In a subsequent hybridization

step the remaining clones were probed with the labelled

recombinant chromosome YHV303-A6 after enrichment by

5   preparative pulsed-field gel electrophoresis.

DNA Sequencing of BAC F15

       DNA of BAC F15 was isolated by an alkaline lysis large

scale plasmid preparation according to Sambrook et al. (1989).

10   50 μg of purified DNA were nebulized by high pressure treatment

with argon gas in a reaction chamber for 150 seconds. The ends

of the sheared and reprecipitated DNA were blunt-ended by a T4

DNA polymerase-mediated fill in reaction. DNA fragments in the

size range between 800 bp and 3 kb were isolated from agarose

15   gels using a DNA isolation kit (Jetsorb, Genomed Inc., U.S.A.),

subcloned into the pBluescript SK vector (Stratagene) and

propagated in *E. coli* DN5α. Clones carrying BAC F15 derived

inserts were selected by hybridization using the sheared DNA of

BAC F15 as a probe. Sequencing reactions were performed as

20   described above. Evaluation of the sequencing data,

construction of sequence contigs, and estimation of coding

propabilities were done by means of the STADEN software package

for Unix users (4th edition, 1994). Assessment of coding

probabilities was based on a combined evaluation of uneven

25   positional base frequencies, positional base preference and

barley codon usage in the investigated contigs. Homology

searches were done using the BLAST software.

89

PCR-based Sequencing of Alleles at Mlo

Plant chromosomal DNA for this purpose was isolated
according to Chunwongse et al. (1993). DNA sequences of *Mlo*
alleles of the different barley varieties, *mlo* mutants, BC
5   lines, and intragenic recombinants used in this study were
obtained by PCR-based sequencing. Seven overlapping
subfragments of the gene (each 400 bp-600 bp in length) were
amplified by PCR (35 cycles, 60°C annealing temperature) using
sets of specific primers. After preparative agarose gel
10   electrophoresis and isolation of the amplification products
using the Jetsorb kit (Genomed Inc., U.S.A.) fragments were
reamplified to increase specificity. The resulting products
were subsequently purified from nucleotides and
oligonucleotides (Jetpure, Genomed Inc., U.S.A.) and used as a
15   template in DNA sequencing reactions (see above). All DNA
sequences of mutant alleles and corresponding regions of the
parental lines and the intragenic recombinants were derived
from both strands and confirmed two times in independent sets
of experiments. In addition, mutant alleles *mlo-1*, *mlo-3*, *mlo-*
20   *4*, *mlo-5*, *mlo-7*, *mlo-8*, *mlo-9*, and *mlo-10* were also verified in
the corresponding BC lines in cultivar Ingrid.

RT-PCR and Rapid Amplification of cDNA Ends (RACE)

RT-PCR was performed using the SUPERSCRIPT
25   preamplification system for first strand cDNA synthesis (Gibco
BRL). Total RNA (1 μg) of seven-day-old primary barley leaves
(cultivar Ingrid) served as template. First strand cDNA

90

synthesis was primed by an oligo(dT) primer. The putative

coding region of the *Mlo* gene was subsequently amplified using

oligonucleotides 25L (GTGCATCTGCGTGTGCGTA) and 38

(CAGAAACTTGTCTCATCCCTG) in a single amplification step (35

5   cycles, 60°C annealing temperature). The resulting product was

analyzed by direct sequencing. 5'- and 3'-ends of the *Mlo* cDNA

were determined by RACE (Frohman et al., 1988) using the

MARATHON cDNA amplification kit (Clontech). Corresponding

experimental procedures were mainly carried out according to

10   the instructions of the manufacturer. To obtain specific RACE

products, two consecutive rounds of amplification (35 cycles,

55°C annealing temperature) were necessary. For this purpose,

two sets of nested primers were used in combination with the

adapter primers of the kit: oligonucleotides 46

15   (AGGGTCAGGATCGCCAC) and 55 (TTGTGGAGGCCGTGTTCC) for the 5'-end

and primers 33 (TGCAGCTATATGACCTTCCCCCTC) and 37

(GGACATGCTGATGGCTCAGA) for the 3'-end. RACE products were

subcloned into pBluescript SK (Stratagene). Ten 5'-end and

eight 3' end clones were chosen for DNA sequence analysis.

20

The term "AFLPs" is used herein to refer to "AFLP

markers".

Table 1 summarizes the identified mutation sites of various mutants within the *Mlo* gene. The origin, the mutagen and the predicted effect of the mutation at the amino acid level are indicated.

Table 2 shows the results of heteroallelic *mlo* crosses and selfings of the respective *mlo* lines to isolate intragenic recombination events.

Table 3 summarizes the genotypes at flanking RFLP markers in susceptible F_2 or homozygous F_3 progeny from the intermutant crosses. CO and NCO indicate crossover type and non crossover type recombinants deduced from flanking molecular marker exchange. Table 3 summarizes DNA sequence analysis of suceptible intragenic crossover type recombinants (from homozygous susceptible F_3 progeny) and the corresponding parental *mlo* mutant lines. Sequences flanking the identified mutation sites are shown.

Table 4 shows the results of direct PCR sequencing of genomic DNA of susceptible intragenic recombinants derived from both heteroallelic crosses *mlo-1 x mlo-8* and *mlo-1 x mlo-5*, revealing restoration of wild type sequences.

Table 5 shows several *Arabidopsis thaliana* and two rice expressed sequence tags (ESTs) with homology to the Mlo protein.

Table 5A show amino acid sequences, with "query" indicating part of the Mlo protein sequence to which homology has been found, with the predicted amino acid sequence of each identified EST marked with "subject".

92

Table 5B shows EST nucleotide sequences encoding the amino
acid sequences shown in Table 5A. GenBank Accession number
T22145 (definition 4153 Arabidopsis thaliana cDNA clone 97N8T7,
NCBI Seq ID 932185), number T22146 (definition 4153 Arabidopsis

5   thaliana cDNA clone 97N9T7, NCBI Seq ID 932186), number N37544
(definition 18771 Arabidopsis thaliana cDNA clone 205N12T7,
NCBI Seq ID 1158686), number T88073 (definition 11769
Arabidopsis thaliana cDNA clone 155I23T7, NCBI Seq ID 935932)
number H76041 (definition 17746 Arabidopsis thaliana cDNA clone

10   193P6T7, NCBI seq ID 1053292), number D24287 (rice cDNA partial
sequence R1638_1A, nID g428139) and D24131 (rice cDNA partial
sequence R1408_1A, nID g427985) are shown. The Arabidopsis
sequences are from Newman et al. (1994) *Plant Physiol.* **106**
1241-55. The rice sequences are from Minobe, Y. and Sasaki, T.

15   submitted 2 Nov 1993 to DDBJ.

Table 1. *mlo* Mutant Alleles

Allele	Mother Variety	Mutagen	Mutational Event at *Mlo*	Effect on Amino Acid Level
mlo-1	Haisa	X-rays	$T^{484} \rightarrow A$	$trp^{162} \rightarrow arg$
mlo-3	Malteria Heda	γ-rays	deletion of 2 nucleotides (1188-1189)	frame shift after phe^{395}
mlo-4	Foma	X-rays	deletion of 11 nucleotides (478-488)	frame shift after trp^{159}
mlo-5	Carlsberg II	EMS	$G^{3} \rightarrow A$	$met^{1} \rightarrow ile^{a}$
mlo-7	Carlsberg II	EMS	$G^{677} \rightarrow A$	$gly^{226} \rightarrow asp$
mlo-8	Carlsberg II	EMS	$A^{1} \rightarrow G$	$met^{1} \rightarrow val^{a}$
mlo-9	Diamant	EMS	$C^{28} \rightarrow T$	$arg^{10} \rightarrow trp$
mlo-10	Foma	γ-rays	deletion of 6 nucleotides (543-548)	2 amino acids (phe^{162}, thr^{163}) missing
mlo-12	Elgina	NMU	$C^{720} \rightarrow A$	$phe^{240} \rightarrow leu$
mlo-13	Plena	EMS	$T^{89} \rightarrow A$	$val^{30} \rightarrow glu$
mlo-16	Alsa	EMS	$G^{1917'} \rightarrow A$	alteration in 3' splice border of intron 9
mlo-17	Plena	EMS	$C^{92} \rightarrow T$	$ser^{31} \rightarrow phe$
mlo-26	Plena	EMS	$T^{809} \rightarrow A$	$leu^{270} \rightarrow his$
mlo-28	Nadja	NaN₃	$C^{665} \rightarrow T$	$thr^{222} \rightarrow ile$

Numbers of nucleotides and amino acids are given according to the translational start site of the *Mlo* cDNA sequence.

Nucleotide number according to the translational start site of the genomic *Mlo* DNA sequence.

EMS = ethylmethane sulfonate, NMU = nitrosomethylurea, NaN₃ = sodium azide.

[a] Next start codon is at nucleotide positions 79-81 and is in frame with the coding sequence.

Table 2

F_2 progeny from intermutant crosses and selfings

Testcrosses	resistant	susceptible	frequency of susceptible F_2 progeny
mlo-8 x mlo-1	5,281	3	5.7×10^{-4}
mlo-5 x mlo-1	915	0	---
mlo-5 x mlo-1	14,474	9	6.2×10^{-4}

selfings	resistant	susceptible	
mlo-1	12,634	0	
mlo-5	5,498	0	
mlo-8	8,435	0	

Table 3. Genotypes at Flanking RFLP Markers in Susceptible Progeny Derived from Heteroallelic mlo Crosses

Testcross	Susceptible Plant	Parental Genotype in Centromeric Orientation to Mlo[1]	Parental Genotype in Telomeric Orientation to Mlo[2]	Type of Recombination
mlo-8 x mlo-1	1	mlo-1	mlo-8	CO
	2	mlo-1	mlo-8	CO
	3	mlo-8	mlo-8	NCO
mlo-1 x mlo-5	1	mlo-1	mlo-5	CO
	2	mlo-1	mlo-5	CO
	3	mlo-5	mlo-5	NCO
	4	mlo-1	mlo-5	CO
	5	mlo-5	mlo-5	NCO
	6	mlo-5	mlo-5	NCO
	7	mlo-1	mlo-5	CO
	8*	mlo-5	mlo-5	NCO
	9*	mlo-1 + mlo-5	mlo-5	CO

[1] deduced from alleles of RFLP marker WG114 (see Figure 1)

[2] deduced from alleles of RFLP marker ABG366 (see Figure 1)

CO = cross over type, NCO = non cross over type of recombination

*Genotypes of flanking RFLP markers have been determined in heterozygous susceptible F_2 individuals; in all other cases homozygous susceptible F_3 progeny derived from the susceptible F_2 individuals were tested

9.6

Table 4. Restoration of *Mlo* Wild Type Sequences by Intragenic Recombinantion Events

Genotypes		Nucleotide Sequences Flanking Mutant Sites[1]	
		Nucleotides -3 to +3	Nucleotides 481 to 486
Haisa	*Mlo*	CCGATG	AATGGG
	mlo-1	CCGATG	AA<u>A</u>GGG
Carlsberg II	*Mlo*	CCGATG	AATGGG
	mlo-5	CCGAT<u>A</u>	AATGGG
	mlo-8	CCG<u>G</u>TG	AATGGG
Intragenic recombinant	*mlo-1 x mlo-8* 1	CCGATG	AATGGG
	2	CCGATG	AATGGG
	mlo-1 x mlo-5 1	CCGATG	AATGGG
	2	CCGATG	AATGGG
	4	CCGATG	AATGGG
	7	CCGATG	AATGGG
	9	CCGATG	AATGGG

[1]Numbers of nucleotides are given according to the translational start site (see Figure 2)

TABLE 5A

>EM EST1:AT1452 T22145 4153 Arabidopsis thaliana cDNA clone 97N8T7. 11/95
  Length = 382

   Plus Strand HSPs:

  Score = 248 (115.9 bits), Expect = 2.9e-27, P = 2.9e-27
  Identities = 47/100 (47%), Positives = 67/100 (67%), Frame = +2

Query: 242 KYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVGTKLEMI 301
              KY+ R++EDDFK VVGIS LW ++   L++NG T   WI+FIP +LL VGTKLE +
Sbjct: 2 KYMMRALEDDFKQVVGISWYLWXFVVIFXLLNVNGWHTYFWIAFIPFXLLLAVGTKLEHV 181

Query: 302 IMEMALEIQDRASVIKGAPVVEPSNKFFWFHRPDWVLFFI 341
              I ++A E+ ++   I+G VV+P . + FWF +P VL+ I
Sbjct: 182 IAQLAHEVAEKHVAIEGDLVVKPXXEHFWFSKPQIVLYLI 301

>EM EST1:AT1462 T22146 4154 Arabidopsis thaliana cDNA clone 97N9T7. 11/95
  Length = 390

   Plus Strand HSPs:

  Score = 212 (99.1 bits), Expect = 4.2e-26, Sum P(2) = 4.2e-26
  Identities = 41/83 (49%), Positives = 58/83 (69%), Frame = +2

Query: 242 KYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVGTKLEMI 301
              KY+ R++EDDFK VVGIS LW   ++ L L++NG T   WI+FIP +LL VGTKLE +
Sbjct: 2 KYMMRALEDDFKQVVGISWYLWXFVVIFLLLNVNGWHTYFWIAFIPFALLLAVGTKLEHV 181

Query: 302 IMEMALEIQDRASVIKGAPVVEP 324
              I ++A E+ ++   I+G VV+P
Sbjct: 182 IAQLAHEVAEKHVAIEGDLVVKP 250

  Score = 52 (24.3 bits), Expect = 1.9, Sum P(2) = 0.85
  Identities = 9/32 (28%), Positives = 16/32 (50%), Frame = +2

Query: 18 WAVAVVFAAMVLVSVLMEHGLHKLGHWFQHRH 49
              W + FA ++ V +EH + +L H   +H
Sbjct: 122 WIAFIPFALLLAVGTKLEHVIAQLAHEVAEKH 217

  Score = 49 (22.9 bits), Expect = 4.2e-26, Sum P(2) = 4.2e-26
  Identities = 8/17 (47%), Positives = 12/17 (70%), Frame = +1

Query: 323 EPSNKFFWFHRPDWVLF 339
              E S++ FWF +P VL+
Sbjct: 244 ETSDEHFWFSKPQXVLY 294

98

TABLE 5A cont'd

>EM EST1:AT54418 N37544 18771 Arabidopsis thaliana cDNA clone 20SN12T7. 1/96
   Length = 585

   Plus Strand HSPs:

   Score = 277 (129.5 bits), Expect = 1.2e-45, Sum P(2) = 1.2e-45
   Identities = 51/96 (53%), Positives = 71/96 (73%), Frame = +1

```
Query:   236 SKFDFHKYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPLVILLCVG 295
             S+FDF KYI+RS+E DFK VV.IS +W VA+L L + G+ + +W+ FIPLV++L VG
Sbjct:   127 SRFDFRKYIQRSLEKDFKTVVEISPVIWFVAVLFLLTNSYGLRSYLWLPFIPLVVILIVG 306

Query:   296 TKLEMIIMEMALEIQDRASVIKGAPVVEPSNKFFWF 331
             TKLE+II ++ L IQ+ V++GAPVV+P + FWF
Sbjct:   307 TKLEVIITKLGLRIQEEGDVVRGAPVVQPGDDXFWF 414
```

   Score = 121 (56.6 bits), Expect = 1.2e-45, Sum P(2) = 1.2e-45
   Identities = 25/45 (55%), Positives = 29/45 (64%), Frame = +1

```
Query:   196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSQNSKFDF 240.
             S T   W+V.FFRQFF SVTKVDYL L GFI AH + ++ F
Sbjct:     1 SKTRVTLWIVCFFRQFFGSVTKVDYLALXHGFIMAHFAPGNESRF 135
```

>EM EST1:AT04117 H76041 17746 Arabidopsis thaliana cDNA clone 193P6T7. 11/95
   Length = 476

   Plus Strand HSPs:
   Score = 210 (98.2 bits), Expect = 9.0e-36, Sum P(2) = 9.0e-36
   Identities = 43/86 (50%), Positives = 58/86 (67%), Frame = +1

```
Query:   196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSQNSKFDFHKYIKRSMEDDFKVV 255
             ++TP   V FFRQFF SV + DYLTLR GF +AHL+   KF+F +YIK S+EDDFK+V
Sbjct:   124 TTTPFXFNVGCFFRQFFVSVERTDYLTLRHGFXSAHLAPGRKFNFQRYIKXSLEDDFKLV 303

Query:   256 VGISLPLWGVAILTLFLDINGVGTLI 281
             VGI LW ++ L + +GT++
Sbjct:   304 VGIXPVLWASFVIFLAVQX*WLGTIV 381
```

   Score = 119 (55.6 bits), Expect = 9.0e-36, Sum P(2) = 9.0e-36
   Identities = 24/57 (42%), Positives = 32/57 (56%), Frame = +1

```
Query:   156 MRTWKKWETETTSLEYQFANDPARFRFTHQTSFVKRHLGLSSTPGIRWVVAFFRQFF 212
             +R WKKWE T S +Y F D +R R TH+TSFV+ H    +T   + V F + F.
Sbjct:     1 IRGWKKWEQXTLSNDYXFXIDHSRLRLTHETSFVREHTSFWTTTPFXFNVGCFFRQF 171
```

   Score = 40 (18.7 bits), Expect = 1.2e-08, Sum P(2) = 1.2e-08
   Identities = 8/19 (42%), Positives = 10/19 (52%), Frame = +2
```
Query:   269 TLFLDINGVGTLIWISFIP 287
             +L + NG G L W S P
Sbjct:   344 SLLFNXNGWGPLFWASVPP 400
```

99

TABLE 5A cont'd

>EM EST1:AT0739 T88073 11769 Arabidopsis thaliana cDNA clone 155I23T7. 11/95
   Length = 460

   Plus Strand HSPs:

   Score = 175 (81.8 bits), Expect = 1.2e-24, Sum P(2) = 1.2e-24
   Identities = 31/67 (46%), Positives = 43/67 (64%), Frame = +1

Query: 146 VITIALSRLKMRTWKKWETETTSLEYQFANDPARFRFTHQTSFVKRHLGLSSTPGIRWVV 205
             ++T A ++KMRTWK WE ET ++EYQ++NDP RFRF   TSF +RHL S +   +
Sbjct: 4 IVTYAFGKIKMRTWKSWEEETKTIEYQYSNDPERFRFARDTSFGRRHLNFWSKTRVTLWI 183

   Score = 121 (56.6 bits), Expect = 1.4e-14, Sum P(2) = 1.4e-14
   Identities = 25/45 (55%), Positives = 29/45 (64%), Frame = +1

Query: 196 SSTPGIRWVVAFFRQFFRSVTKVDYLTLRAGFINAHLSQNSKFDF 240
             S T   W+V FFRQFF SVTKVDYL L GFI AH + ++ F
Sbjct: 157 SKTRVTLWIVCFFRQFFGSVTKVDYLALXHGFIMAHFAPGNESRF 291

   Score = 75 (35.1 bits), Expect = 1.2e-24, Sum P(2) = 1.2e-24
   Identities = 14/21 (66%), Positives = 17/21 (80%), Frame = +1

Query: 236 SKFDFHKYIKRSMEDDFKVVV 256
             S+FDF KYI+RS+ DFK VV
Sbjct: 283 SRFDFRKYIQRSLXXDFKTVV 345

>EM EST5:OSR16381A D24287 Rice cDNA, partial sequence (R1638_1A). 5/95
   Length = 400
   Plus Strand HSPs:

   Score = 147 (68.7 bits), Expect = 1.9e-16, Sum P(2) = 1.9e-16
   Identities = 26/53 (49%), Positives = 35/53 (66%), Frame = +1

Query: 236 SKFDFHKYIKRSMEDDFKVVVGISLPLWGVAILTLFLDINGVGTLIWISFIPL 288
             ++F+F KYIKR +EDDFK VVGIS P W A+ + +++G L W S PL
Sbjct: 202 TRFNFRKYIKRXLEDDFKTVVGISAPXWASALAIMLFNVHGWHNLFWFSTXPL 360

   Score = 45 (21.0 bits), Expect = 1.9e-16, Sum P(2) = 1.9e-16
   Identities = 9/15 (60%), Positives = 11/15 (73%), Frame = +2

Query: 287 PLVILLCVGTKLEMI 301
             PL + L VGTKL+ I
Sbjct: 356 PLXVTLAVGTKLQAI 400

>EM EST5:OSS1692A D39989 Rice cDNA, partial sequence (S1692_1A). 11/94
   Length = 343

   Plus Strand HSPs:
   Score = 95 (44.4 bits), Expect = 0.00059, P = 0.00059
   Identities = 24/58 (41%), Positives = 31/58 (53%), Frame = +3

Qu ry: 43 HWFQHRHKKALWEALEKMKAELMLVGFISLLLIVTQDPIIAKICISEDAADVMWPCKR 100
             H + H+ L +A+EKMK E++ML+GFISLLL T I   S+   PC R
Sbjct: 3 HXSEKTHRNPLHKAMEKMKEEMMLLGFISLLLAATSRIISGICIDSKYYNSNFSPCTR 176

TABLE 5B   100

GenBank Accession Number T22145


```
  1 caagtatatg atgcgcgctc tagaggatga tttcaaacaa gttgttggta ttagttggta
 61 tctttggntc tttgtcgtca tcttttttnct gctaaatgtt aacggatggc acacatattt
121 ctggatagca tttattccct ttnctttgct tcttgctgtg ggaacaaagt tggagcatgt
181 nattgcacag ttagctcatg aagttgcaga gaaacatgta gccattgaag gagacttagt
241 ggtgaaaccc ncanatgagc atttctggtt cagcaaacct caaattgttc tctacttgat
301 cccattttat cctctttccc agaatgcntt ttnagantgc nttttttnnt tttggnnttt
361 ggggtaanan annggtttcg nc
```


GenBank Accession Number T22146


```
  1 caagtatatg atgcgcgctc tagaggatga tttcaaacaa gttgttggta ttagttggta
 61 tctttggntc tttgtcgtca tcttttttgct gctaaatgtt aacggatggc acacatattt
121 ctggatagca tttattccct ttgctttgct tcttgctgtg ggaacaaagt tggagcatgt
181 nattgcacag ttagctcatg aagttgcaga gaaacatgta gccattgaag gagacttagt
241 ggtgaaacct cagatgagca tttctggttc agcaaacctc aaantgttct ctactngatc
301 cnctttatcc cccttccaga atgccttttt nangattcnn nttttttcctt nttgganntt
361 ttgggnnnnc aaacgggntt nggacctccg
```


GenBank Accession Number N37544


```
  1 agcaagacga gagtcacact atggattgtt tgttttttta gacagttctt tggatctgtc
 61 accaaagttg attacttagc actaagncat ggtttcatca tggcgcattt tgctcccggt
121 aacgaatcaa gattcgattt ccgcaagtat attcagagat cattagagaa agacttcaaa
181 accgttgttg aaatcagtcc ggttatctgg tttgtcgctg tgctattcct cttgaccaat
241 tcatatggat tacgttctta cctctggtta ccattcattc cactagtcgt aattctaata
301 gttggaacaa agcttgaagt cataataaca aaattgggtc taaggatcca agaggaaggt
361 gatgtggtga gaggcgcccc agtggttcag cctggtgatg accncttctg gtttngnaan
421 cacgnttcaa tnttttccnt antcacttng gccttttttan gggtgaattt caacttcatn
481 ctttncctgg ggncggatga ttcaatccaa naatnttccc ctgaagnctn caagtttggg
541 cataggcttt nggtgggntt ttcaganttt nagtttggct tnccc
```


SUBSTITUTE SHEET (RULE 26)

TABLE 5B (Continued)

GenBank Accession Number T88073

```
  1 tgcattgtta cttatgcttt cggaaagatc aagatgagga cgtggaagtc gtgggaggaa
 61 gagacaaaga caatagagta tcagtattcc aacgatcctg agaggttcag gtttgcnagg
121 gacacatctt ttgggagaag acatctcaat ttctggagca agacgagagt cacactatgg
181 attgtttgtt tttttagaca gttctttgga tctgtcacca aagttgatta cttagcacta
241 agncatggtt tcatcatggc gcattttgct cccggtaacg aatcaagatt cgatttccgc
301 aagtatattc agagatcatt agngnaagac ttcaaaaccg ttgtttgaaa tcagtccggt
361 tatctggttt gtcggctgtg ctattccnct tgaccaattc atatggntnc ggtnttncnc
421 tggtaccatt attcnctagc ggaatntaaa agttggcnga
```

GenBank Accession Number H76041

```
  1 attcgtggat ggaaaaagtg ggagcaagan acattatcta atgactatna gtttnctatt
 61 gatcattcaa gacttaggct cactcatgag acttcttttg tnagagaaca tacaagtttc
121 tggacaacaa cncctttctn ctttaacgtc ggatgcttct ttaggcagtt ctttgtatct
181 gtngaaagaa ccgactactt gactctgcgc catggattca nctctgccca tttagctcca
241 ggaagaaagt tcaacttcca gagatatatc aaangatctc tcgaggatga tttcaagttg
301 gtagttggaa taagnccagt tctttgggca tcatttgtaa tcttccttgc tgttcaatgn
361 taatggctgg ggaccattgt tttgggcntc ggtaccgcct ntactcanaa ncccaggctt
421 ttggccaagg ttcaaggaat ttngggacaa tggggtagaa tcgtgggcnc atnngg
```

GenBank Accession Number D24287

```
  1 tcntnttttnn ttttcgnntn cntccacccc tnnnntnctc nancncnttn nnnttatctc
 61 tnttnttntc ncntntcccn ncaccaccnn ncgacgggcn tggactnngc ccnnngttcg
121 aggctgccca ctgncgtctg agacctacct tgncatttga cggcacngga cttcanttgc
181 tgctcacttt atctctacgg gactaggttc aattttcgga aatacatcaa aaggncactg
241 gaggacgatt ttaagacagt tgttggcatt agtgcacccn tatgggcttc tgcgttggcc
301 attatgctct tcaatgttca tggatggcat aacttgttct ggttctctac aatncccctt
361 gntagtaact ttagcagttg gaacaaagct gcaggctata
```

GenBank Accession Number D24131

```
  1 cagactacct gactttgagg cacggattca ttgctgctca tttatctcta gggactaggt
 61 tcaattttcg gaaatacatc aaaaggtcac tggaggacga ttttaagaca gttgttggca
121 ttagtgcacc cttatgggct tctgcgttgg ccattatgct cttnaatgtt catggatggc
181 ataacttgtt ctggttctct acaatccccc ttgtagtaac tttagcagtt ggaacaaagc
241 tgcaggctat aaattgcaatg atggctgttg aaattaaaga gaggcataca gtaattcaag
301 gaatgccggt ggtgaactca gtgat
```

102

REFERENCES

1.   Bayles, et al. (1990). Physiol. Mol. Plant Pathol. 36,
     63-72.

2.   Becker, et al. (1995). Mol. and Gen. Genet. 249, 65-73.

3.   Bennet, et al. (1991). Phil. Trans. R. Soc. London (Biol)
     334, 309-345.

4.   Bent, et al. (1994). Science 265, 1856-1860.

5.   Boyd, et al. (1995). Plant. J., 7, 959-968.

6.   Bourque, J. E. (1995). Plant Science 105, 125-149.

7.   Burke, et al. (1987). Science 236, 806-812.

8.   Cao, et al. (1992). Plant Cell Rep. 11, 585-591.

9.   Christou, et al. (1991). Bio/Technology 9, 957-962.

10. Chunwongse, et al. (1993). Ther. Appl. Genet., 86, 694-
     698.

11. Civardi (1994).

12. Cornejo, et al. (1993). Plant Molecular Biology 23, 567-
     581.

13. Dangl, J. L. (1995). Cell 80, 363-366.

14. Dangl, J. L., et al. (1996). Plant Cell (*in press*).

15. Datta, et al. (1990). Bio/Technology 8, 736-740.

16. de Feyter, et al. (1996). Molecular and General Genetics
     250, 329-338.

17. D'Halluin, et al. (1992). Plant Cell 4, 1495-1505.

18. Dietrich, et al. (1994). Cell 77, 565-577.

19. Doke, N. (1983). Physiol. Plant Pathol., 23, 345-357.

20. Doke, N. and Ohashi, Y. (1988). Physiol. Mol. Plant
     Path., 32, 163-175.

103

21. Dooner, H.K. and Kermicle, J.L. (1986). Genetics, 113, 135-143.

22. Dujon, B. (1996). Trends Genet., 12, 263-270.

23. Flavell, R. B. (1994). Proc. Natl. Acad. Sci. USA 91, 3490-3496.

24. Flor, H. H. (1971). Annu. Rev. Phytopathol. 9, 275-296.

25. Freeling, M. (1978). Genetics, 89, 211-224.

26. Freialdenhoven, A., et al. (1996). Plant Cell, 8, 5-14.

27. Frohman, N.A., et al. (1988). Proc. Natl. Acad. Sci., 85, 8998-9002.

28. Fromm, et al. (1990). Bio/Technology 8, 833-839.

29. Fugimoto, et al. (1993). Bio/Technology 11, 1151-1155.

30. Giovannoni, J.J., et al. (1991). Nucl. Acids Res., 19, 6553-6558.

31. Gordon-Kamm, et al. (1990). Plant Cell 2, 603-618.

32. Görg, R., et al. (1993). Plant J., 3, 857-866.

33. Grant, et al. (1995). Science 269, 843-846.

34. Greenberg, et al. (1993). The Plant Journal 4, 327-341.

35. Greenberg, et al. (1994). Cell 77, 551-563.

36. Habekuss, A. and Hentrich, W. (1988). *Tag. Ber., Akad. Lanwirtsch.-Wiss. DDR, Berlin*, 272: 229-237.

37. Hammond-Kosack, K.E. and Jones, J.D.G. (1996). *The Plant Cell (in press)*.

38. Hartmann, E., *et al.* (1989). *Proc. Natl. Acad. Sci.*, 86: 5786-5790.

39. Haselhoff, et al. (1988). Nature 334, 585-491.

40. Heath, M.C. (1980). *Annu. Rev. Phytopathol.*, 18: 211-236.

104

41. Hentrich, W. (1979).   Arch. Züchtungsvorsch., Berlin 9, S. 283-291.

42. Heslop-Harrison, J.S. (1991). J. Cell Sci., 100, 15-21.

43. Hiei, et al. (1994). The Plant Journal 6, 271-282.

44. Hieter, et al. (1990). Cold Spring Harbor, NY, Col Spring Harbor Press.

45. Hinze, et al. (1991). Proc. Natl. Acad. Sci. USA 88, 3691-3695.

46. Jabs, T., et al. (1996). Science, 273, 1853-55.

47. Jones, D.T., et al. (1994). Biochemistry, 33, 3038-3049.

48. Jones, et al. (1994). Science 266, 789-793.

49. Jones, J. D. G. (1994). Current Biology 4, 749-751.

50. Jørgensen, J. H. (1977). Induced mutations against plant diseases (Crop Symposium Vienna)533-547.

51. Jørgensen, J. H. (1983). Induced mutations for disease resistance in crop plants II (International Atomic Energy Agency, Vienna)73-87.

52. Jørgensen, J. H. (1992). Euphytica 63, 141-152.

53. Jørgensen, J. H. (1994). Critical Reviews in Plant Sciences 13, 97-119.

54. Jørgensen, J.H. and Mortensen, K. (1977). Phytopathology, 67, 678-685.

55. Klein, P., et al. (1985). Biochim. Biophys. Acta, 815, 468-476.

56. Koga, H., et al. (1990). Can. J. Bot., 68, 2344-2352.

57. Koornneef, M., et al. (1983). Genet. Res. Camb., 41, 57-68.

58.   Kosslak, R., et al. (1996). J. Hered (*in press*).

59.   Kuhn, R.M. and Ludwig. R.A. (1994). Gene, 141, 125-127.

60.   Künzel, G. (1982). Theor. Appl. Genet., 64, 25-29.

61.   Koziel, et al. (1993). Biotechnology 11, 194-200.

62.   Lamb, C.J. (1994). *Cell*, 76: 419-422.

63.   Lawrence, et al. (1995). The Plant Cell 7, 1195-1206.

64.   Levine, A., *et al.* (1994). *Cell*, 79: 583-593.

65.   Li, et al. (1993). Plant Cell Rep. 12, 250-255.

66.   Linde-Laursen, I., et al. (1982). Z. Pflanzenzüchtg., 81,
      191-219.

67.   Lundqvist, U., et al. (1991). Hereditas, 115, 227-239.

68.   Martin, et al. (1993). Science 262, 1432-1436.

69.   McClintock, B. (1984). Science, 226, 792-801.

70.   Mindrinos, et al. (1994). Cell 78, 1089-1099.

71.   Moore, et al. (1995). Current Biology 5, 737-739.

72.   Mourad, G., et al. (1994). Mol. Gen. Genet., 243, 178-
      184.

73.   Negassa, M. (1985). Hereditas, 102, 113-121.

74.   Nigg, et al. (1991). Cell 66, 15-22.

75.   Osborne, et al. (1995). Current Opinion in Cell Biology
      7, 406-413.

76.   Olsen, O., et al. (1993). Proc. Natl. Acad. Sci. USA, 90,
      8043-8047.

77.   Peng, et al. (1992). Theor Appl Genet 83, 855-863.

78.   Peng, et al. (1991). International Rice Research
      Institute, Manila, Philippines.563-574.

106

79. Pryor, A. J. (1987). Trends in Genetics 3, 157-161.

80. Raff, M.C. (1992). Nature, 356, 397-400.

81. Raff, M.C., et al. (1993). Science, 262, 695-700.

82. Rathore, et al. (1993). Plant Molecular Biology 21, 871-
     884.

83. Rihs, et al. (1991). EMBO J. 10, 633-639.

84. Ryerson, D.E. and Heath, M.C. (1996). Plant Cell, 8, 393-
     402.

85. Salamini, F., and Lorenzoni, C. (1970). Mol. Gen. Genet.,
     108, 225-232.

86. Sambrook, J., et al. (1989). Cold Spring Harbor
     Laboratory Press, New York.

87. Schmidt, R., et al. (1995). Science, 270, 480-483.

88. Sherman, et al. (1986). Cold Spring Harbor, NY, Cold
     Spring Harbor Laboratory Press.

89. Shimamoto, K. (1994). Current Opinion in Biotechnology 5,
     158-162.

90. Shimamoto, et al. (1989). Nature 338, 274-276.

91. Shirley, B.W., et al. (1992). Plant Cell, 4, 333-347.

92. Shizuya, et al. (1992). Proc. Natl. Acad. Sci. USA 89,
     8794-8797.

93. Siedler, et al. (1991). Molecular and General Genetics
     226, 117-181.

94. Søgaard, et al. (1988). Barley Genet. Newslett. 17, 120-
     134.

95. Somers, et al. (1992). Bio/Technology 10, 1589-1594.

96. Song, et al. (1995). Science 270, 1804-1806.

107

97. Stakman, E.C. (1915). *J. Ag. Research*, 4: 193-199.

98. Staskawicz, et al. (1995). Science 268, 661-667.

99. Stewart, C.N. and Via, L.E. (1993). *BioTechniques*, 14: 748-750.

5   100. Szostak, J.W., *et al.* (1983). *Cell*, 33: 25-35.

101. Tanksley, et al. (1995). Trends in Genetics 11, 63-68.

102. Thomas, C.M., et al. (1995). *Plant J.*, 8: 785-794.

103. Toriyama, et al. (1988). Bio/Technology 6, 1072-1074.

104. Tsuji, et al. (1992). Plant Physiology 98, 1304-1309.

10   105. Vasil, I. K. (1994). Plant Molecular Biology 25, 925-937.

106. Vasil, et al. (1992). Bio/Technology 10, 667-674.

107. Vos, et al. (1995). Nucleic Acids Research 23, 4407-4414.

108. Walbot, et al. (1983). New York, Plenum Press431-442.

109. Walters, et al. (1992). Plant Molecular Biology 18, 189-

15   200.

110. Wang, et al. (1995). Plant Journal 7, 525-533.

111. Weeks, et al. (1993). Plant Physiology 102, 1077-1084.

112. Weymann, et al. (1995). The Plant Cell 7, 2013-2022.

113. White, E. (1996). *Genes & Development*, 10: 1-15.

20   114. Whitham, et al. (1994). Cell 78, 1011-1115.

115. Wiberg, A. (1974). Hereditas 77, 89-148.

116. Wolter, et al. (1993). Mol. Gen. Genet. 239, 122-128.

117. Wyllie, A.H. (1995). *Current Biology*, 5: 97-104.

118. Zhang, et al. (1988). Plant Cell Rep. 7, 379-384.

25   119. Zhang, et al. (1991). Plant Cell 3, 1155-1165.

120. Zhang, et al. (1988). Theor Appl Genet 76, 835-840.

108

CLAIMS:

1.   An isolated polynucleotide encoding a polypeptide which
includes the amino acid sequence shown in Figure 2.

5

2.   A polynucleotide according to claim 1 wherein the coding
sequence is the coding sequence shown in Figure 2.

3.   A polynucleotide according to claim 1 wherein the coding

10   sequence is a mutant, allele, variant or derivative of the
coding sequence shown in Figure 2, by way of addition,
deletion, substitution and/or insertion of one or more
nucleotides.

15   4.   An isolated polynucleotide which on expression in a
transgenic plant exerts a negative regulatory effect on a
pathogen defence response of the plant, which defence response
is pathogen independent and autonomous of the presence of
pathogen, the polynucleotide encoding a polypeptide which

20   includes an amino acid sequence which is a mutant, allele,
variant or derivative of the Barley Mlo sequence shown in
Figure 2, or is a homologue of another species or a mutant,
allele, variant or derivative thereof, the amino acid sequence
differing from that shown in Figure 2 by way of addition,

25   substitution, deletion and/or insertion of one or more amino
acids.

109

5.   A polynucleotide according to claim 4 encoding a
polypeptide which includes the amino acid sequence shown in
Figure 13.

5   6.   A polynucleotide according to claim 5 wherein the coding
sequence is that shown in Figure 10.

7.   A polynucleotide according to claim 5 wherein the coding
sequence is a mutant, allele, variant or derivative of the
10   coding sequence shown in Figure 10, by way of addition,
deletion, substitution and/or insertion of one or more
nucleotides.

8.   A polynucleotide according to claim 4 encoding a
15   polypeptide which includes the amino acid sequence shown in
Figure 14.

9.   A polynucleotide according to claim 8 wherein the coding
sequence is that shown in Figure 11.

20

10.   A polynucleotide according to claim 8 wherein the coding
sequence is a mutant, allele, variant or derivative of the
coding sequence shown in Figure 11, by way of addition,
deletion, substitution and/or insertion of one or more
25   nucleotides.

11.   A polynucleotide according to claim 4 encoding a

110

polypeptide which includes the amino acid sequence shown in
Figure 15.

12.   A polynucleotide according to claim 11 wherein the coding

5   sequence is that shown in Figure 12.

13.   A polynucleotide according to claim 11 wherein the coding
sequence is a mutant, allele, variant or derivative of the
coding sequence shown in Figure 12, by way of addition,

10   deletion, substitution and/or insertion of one or more
nucleotides.

14.   A polynucleotide according to any preceding claim operably
linked to a regulatory sequence for expression.

15

15.   An isolated polynucleotide encoding a polypeptide which on
expression in a transgenic plant produces a polypeptide which
can stimulate or maintain a defence response of the plant, the
encoded polypeptide including an amino acid sequence which is a

20   mutant, allele, variant or derivative of the Barley Mlo
sequence shown in Figure 2 or of a homologue of another
species, the amino acid sequence differing from that shown in
Figure 2 by way of addition, substitution, deletion and/or
insertion of one or more amino acids.

25

16.   A polynucleotide according to claim 15 which stimulates or
maintains said defence response of the plant on homozygous

111

expression in the plant.

17.   A polynucleotide according to claim 15 wherein the amino acid sequence includes an alteration identified in Table 1.

18.   A polynucleotide according to claim 17 wherein the amino acid sequence is that of Figure 2 including a substitution at residue 240.

19.   A polynucleotide according to claim 17 wherein the amino acid sequence includes Leucine at residue 240.

20.   A polynucleotide according to any of claims 15 to 19 operably linked to a regulatory sequence for expression.

21.   An isolated polynucleotide which has at least about 600 contiguous nucleotides of the nucleotide sequence of any of claims 1 to 13 or complement thereof

22.   A polynucleotide according to claim 21 operably linked to a regulatory sequence for transcription.

23.   An isolated polynucleotide which has at least about 300 contiguous nucleotides of the sequence of any of claims 1 to 13, or complement thereof, operably linked to a regulatory sequence for transcription.

112

24.   A polynucleotide according to claim 22 or claim 23 wherein
the regulatory sequence includes an inducible promoter.

25.   A nucleic acid vector suitable for transformation of a
host cell and including a polynucleotide according to any
preceding claim.

26.   A nucleic acid vector according to claim 25 wherein said
host cell is a microbial cell.

27.   A nucleic acid vector according to claim 25 wherein said
host cell is a plant cell.

28.   A host cell containing a heterologous polynucleotide or
nucleic acid vector according to any preceding claim.

29.   A cell according to claim 28 which is microbial.

30.   A cell according to claim 28 which is a plant cell.

31.   A cell according to claim 30 having said heterologous
polynucleotide incorporated within its genome.

32.   A cell according to claim 31 having more than one said
polynucleotide per haploid genome.

33.   A cell according to any of claims 30 to 32 which is

comprised in a plant.

34.   A plant including a cell according to any of claims 30 to
32.

5

35.   A plant which is a sexually or asexually propagated off-
spring, clone or descendant of a plant according to claim 34,
or any part or propagule of said plant, off-spring, clone or
descendant.

10

36.   A part or propagule of a plant according to claim 35.

37.   A plant according to claim 34 which does not breed true.

15   38.   A method of producing a plant, the method including
incorporating a heterologous polynucleotide according to any of
claims 1 to 14 into a plant cell and regenerating a plant from
said plant cell.

20   39.   A method of producing a plant, the method including
incorporating a heterologous polynucleotide according to any of
claims 15 to 20 into a plant cell and regenerating a plant from
said plant cell.

25   40.   A method of producing a plant, the method including
incorporating a heterologous polynucleotide according to any of
claims 21 to 24 into a plant cell and regenerating a plant from

said plant cell.

41.   A method according to any of claims 38 to 40 including

sexually or asexually propagating or growing off-spring or a

5   descendant of said plant.

42.   A method of stimulating a defence response in a plant, the

method including causing or allowing transcription from a

heterologous polynucleotide according to any of claims 1 to 14

10   within cells of the plant.

43.   A method of stimulating a defence response in a plant, the

method including causing or allowing transcription from a

heterologous polynucleotide according to any of claims 15 to 20

15   within cells of the plant.

44.   A method of stimulating a defence response in a plant, the

method including causing or allowing transcription from a

heterologous polynucleotide according to any of claims 21 to 24

20   within cells of the plant.

45.   A method of producing a polynucleotide encoding a

polypeptide which on expression in a transgenic plant produces

a polypeptide which can stimulate or maintain a defence

25   response of the plant, the method including alteration of the

nucleotide sequence of a polynucleotide according to any of

claims 1 to 14.

115

46.   A method according to claim 45 involving site-specific
sequence mutation.

47.   A method according to claim 45 involving intracellular
5   homologous recombination.

48.   A method wherein following alteration of a nucleotide
sequence in accordance with the method of claim 45 a
polynucleotide including the altered nucleotide sequence is
10   introduced into a host cell.

49.   A method according to claim 48 wherein the host cell is a
plant cell.

15   50.   A method wherein following introduction of a
polynucleotide into a plant cell in accordance with claim 49 a
plant is regenerated from the cell or descendants thereof
including the altered nucleotide sequence.

20   51.   Use of a polynucleotide according to any of claims 1 to 14
for stimulating a defence response in a plant.

52.   Use of a polynucleotide according to any of claims 15 to
20 for stimulating a defence response in a plant.

25

53.   Use of a polynucleotide according to any of claims 21 to
24 for stimulating a defence response in a plant.

54.   Use of a polynucleotide according to any of claims 21 to
24 for down-regulation of expression of a gene encoded a
polypeptide encoded by a polynucleotide according to any of
claims 1 to 14.

5

55.   Use of a polynucleotide according to any of claims 1 to 14
in the production of a transgenic plant.

56.   Use of a polynucleotide according to any of claims 15 to
10   20 in the production of a transgenic plant.

57.   Use of a polynucleotide according to any of claims 21 to
24 in the production of a transgenic plant.

15   58.   A method of determining the presence of a pathogen
resistance or susceptibility allele in a plant or plant cell,
the method including analysing a sample from the plant or plant
cell by:

      (a) comparing the sequence of nucleic acid in the sample
20   with all or part of the nucleotide sequence shown in Figure 7
to determine whether the sample from the patient contains a
mutation;

      (b) determining the presence in the sample of a
polypeptide including the amino acid sequence shown in Figure 7
25   or a fragment thereof and, if present, determining whether the
polypeptide is full length, and/or is mutated, and/or is
expressed at the normal level;

117

(c) performing DNA fingerprinting to compare the
restriction pattern produced when a restriction enzyme cuts
nucleic acid in the sample with the restriction pattern
obtained from the nucleotide sequence shown in Figure 7 or from
5   a known mutant, allele or variant thereof;

(d) contacting the sample with a specific binding member
capable of binding to nucleic acid including the nucleotide
sequence as set out in Figure 7 or a fragment thereof, or a
mutant, allele or variant thereof, the specific binding member
10   including nucleic acid hybridisable with the sequence of Figure
7 or a polypeptide including a binding domain with specificity
for nucleic acid including the sequence of Figure 7 or the
polypeptide encoded by it, or a mutated form thereof, and
determining binding of the specific binding member;

15   (e) performing PCR involving one or more primers based on
the nucleotide sequence shown in Figure 7 to screen the sample
for nucleic acid including the nucleotide sequence of Figure 7
or a mutant, allele or variant thereof.

20   59.   A method of determining the presence of target nucleic
acid in a plant or plant cell, the method including contacting
a nucleic acid molecule which includes the nucleotide sequence
shown in Figure 7 or an oligonucleotide fragment thereof with
nucleic acid in a sample from the plant or plant cell and
25   assessing hybridisation of said nucleic acid molecule with
nucleic acid in the sample.

60. A method according to claim 59 which involves amplification of nucleic acid to which said nucleic acid molecule hybridises.

5 61. A method according to claim 59 or claim 60 wherein said nucleic acid molecule includes an alteration in sequence compared with the nucleotide sequence shown in Figure 7 or corresponding fragment thereof.

10 62. A method according to claim 61 wherein said alteration is selected from those shown in Table 1.

63. An assay method for identifying a compound able to bind the polypeptide encoded by the polynucleotide of any of claims 15 1 to 14 or any of claims 15 to 20, the method including:
(a) bringing into contact said polypeptide or a fragment thereof, and a test compound; and
(b) determining interaction or binding between said polypeptide or fragment thereof and the test compound.

20

64. An assay method according to claim 63 wherein a compound is identified which is able to bind the polypeptide for which the amino acid sequence is shown in Figure 2.

25 65. An assay method for identifying a compound able to stimulate a defence response in a plant by interaction with the polypeptide encoded by the polynucleotide of any of claims 1 to

14 or any of claims 15 to 20, the method including:

(a)   contacting a plant or plant part with a test compound and determining stimulation of a defence response; and

(b)   bringing into contact said polypeptide or a fragment

5   thereof with a test compound and determining interaction or binding between said polypeptide or a fragment thereof and the test compound;

step (b) being performed with a test compound which tests positive in step (a), or step (a) being performed with a test

10   compound which tests positive in step (b), or steps (a) and (b) being performed in parallel.

66.   An assay method according to claim 65 wherein stimulation of a defence response is determined by monitoring pathogen

15   growth and/or viability on the plant or plant part.

67.   An assay method according to claim 65 or claim 66 wherein a compound is identified which is able to bind the polypeptide for which the amino acid sequence is shown in Figure 2.

20

68.   An assay method according to any of claims 65 to 67 wherein a compound is identified which is able to stimulate resistance to powdery mildew in barley.

25   69.   A method which includes following identification of a compound as being able to stimulate a defence response in a plant in accordance with any of claims 65 to 68 formulation of

the compound, or optionally if the compound is peptidyl nucleic
acid encoding it, into a composition including at least one
additional component.

5   70.   A method which includes following identification of a
compound as being able to stimulate a defence response in a
plant in accordance with any of claims 56 to 58 application of
the compound, or optionally if the compound is peptidyl nucleic
acid encoding it, to a plant.

10

71.   Use of a polypeptide encoded by a polynucleotide according
to any of claims 1 to 14, in screening for compounds able to
stimulate a defence response in a plant.

15   72.   Use of a polypeptide encoded by a polynucleotide according
to any of claims 15 to 20, in screening for compounds able to
stimulate a defence response in a plant.

73.   A compound able to stimulate a defence response in a plant
20   identified by a method according to any of claims 63 to 68.
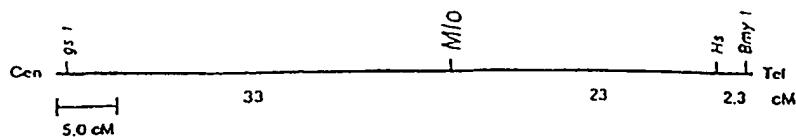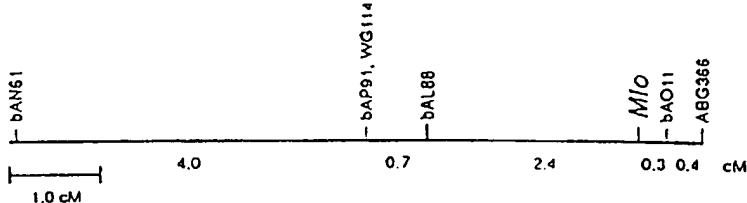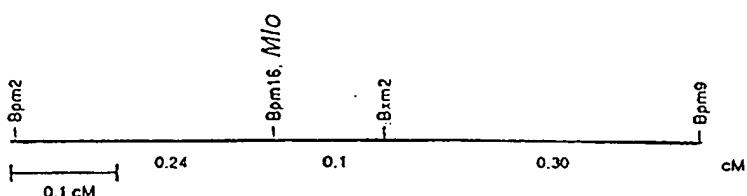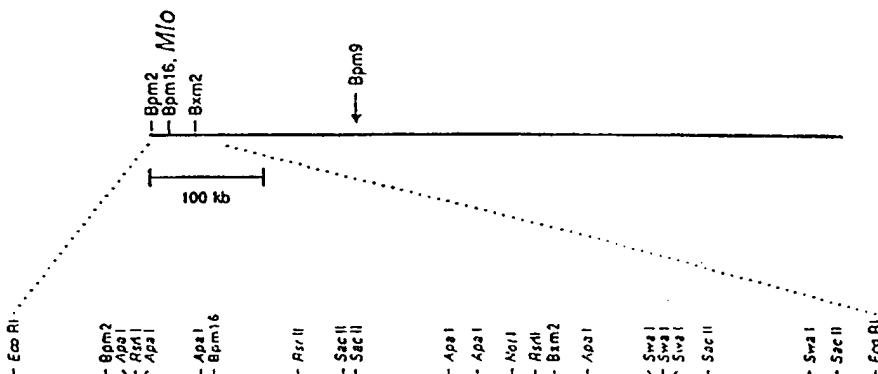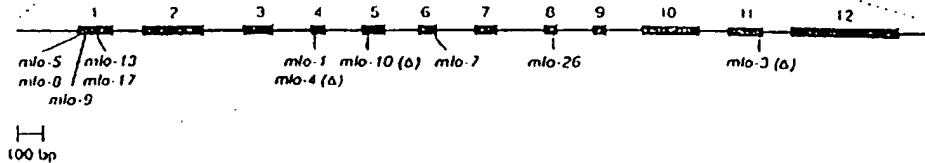
Figure 1

2/28

```
    M   S   D   K   K   G   V   P   A   R   E   L   P   E   T   P   S   W   A   V
ATGTCGGACAAAAAAGGGGTGCCGGCGCGGGAGCTGCCGGAGACGCCGTCGTGGGCGGTG                    60

    A   V   V   F   A   A   M   V   L   V   S   V   L   M   E   H   G   L   H   K
GCGGTGGTCTTCGCCGCCATGGTGCTCGTGTCCGTCCTCATGGAACACGGCCTCCACAAG                    120

    L   G   H   W   F   Q   H   R   H   K   K   A   L   W   E   A   L   E   K   M
CTCGGCCATTGGTTCCAGCACCGGCACAAGAAGGCCCTGTGGGAGGCGCTGGAGAAGATG                    180

    K   A   E   L   M   L   V   G   F   I   S   L   L   L   I   V   T   Q   D   P
AAGGCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCAGGACCCC                    240

    I   I   A   K   I   C   I   S   E   D   A   A   D   V   M   W   P   C   K   R
ATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCATGTGGCCCTGCAAGCGC                    300

    G   T   E   G   R   K   P   S   K   Y   V   D   Y   C   P   E   G   K   V   A
GGCACCGAGGGCCGCAAGCCCAGCAAGTACGTTGACTACTGCCCGGAGGGCAAGGTGGCG                    360

    L   M   S   T   G   S   L   H   Q   L   H   V   F   I   F   V   L   A   V   F
CTCATGTCCACGGGCAGCTTGCACCAGCTGCACGTCTTCATCTTCGTGCTCGCGGTCTTC                    420

    H   V   T   Y   S   V   I   T   I   A   L   S   R   L   K   M   R   T   W   K
CATGTCACCTACAGCGTCATCACCATAGCTCTAAGCCGTCTCAAAATGAGAACATGGAAG                    480

    K   W   E   T   E   T   T   S   L   E   Y   Q   F   A   N   D   P   A   R   F
AAATGGGAGACAGAGACCACCTCCTTGGAATACCAGTTCGCAAATGATCCTGCACGGTTC                    540

    R   F   T   H   Q   T   S   F   V   K   R   H   L   G   L   S   S   T   P   G
CGGTTCACGCACCAGACGTCGTTCGTGAAGCGCCACCTGGGCCTCTCCAGCACCCCTGGC                    600

    I   R   W   V   V   A   F   F   R   Q   F   F   R   S   V   T   K   V   D   Y
ATCAGATGGGTGGTGGCCTTCTTCAGGCAGTTCTTCAGGTCAGTCACCAAGGTGGACTAC                    660

    L   T   L   R   A   G   F   I   N   A   H   L   S   Q   N   S   K   F   D   F
CTGACCTTGAGGGCAGGCTTCATCAACGCGCATTTGTCGCAAAACAGCAAGTTCGACTTC                    720

    H   K   Y   I   K   R   S   M   E   D   D   F   K   V   V   V   G   I   S   L
CACAAGTACATCAAGAGGTCGATGGAGGACGACTTCAAGGTCGTCGTCGGCATCAGCCTC                    780

    P   L   W   G   V   A   I   L   T   L   F   L   D   I   N   G   V   G   T   L
CCGCTGTGGGGTGTGGCGATCCTCACCCTCTTCCTTGACATCAATGGGGTTGGCACGCTC                    840

    I   W   I   S   F   I   P   L   V   I   L   L   C   V   G   T   K   L   E   M
ATCTGGATTTCTTTCATCCCTCTCGTGATCCTCTTGTGTGTTGGAACCAAGCTGGAGATG                    900

    I   I   M   E   M   A   L   E   I   Q   D   R   A   S   V   I   K   G   A   P
ATCATCATGGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCGTCATCAAGGGGGCCCCC                    960

    V   V   E   P   S   N   K   F   F   W   F   H   R   P   D   W   V   L   F   F
GTGGTCGAGCCCAGCAACAAGTTCTTCTGGTTCCACCGCCCCGACTGGGTCCTCTTCTTC                    1020

    I   H   L   T   L   F   Q   N   A   F   Q   M   A   H   F   V   W   T   V   A
ATACACCTGACGTTGTTCCAGAACGCGTTTCAGATGGCGCATTTTGTGTGGACAGTGGCC                    1080

    T   P   G   L   K   K   C   Y   H   T   Q   I   G   L   S   I   M   K   V   V
ACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGGCTGAGCATCATGAAGGTGGTG                    1140

    V   G   L   A   L   Q   F   L   C   S   Y   M   T   F   P   L   Y   A   L   V
GTGGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCTTCCCCCTCTACGCGCTCGTC                    1200

    T   Q   M   G   S   N   M   K   R   S   I   F   D   E   Q   T   S   K   A   L
ACACAGATGGGATCAAACATGAAGAGGTCCATCTTCGACGAGCAGACGTCCAAGGCGCTC                    1260

    T   N   W   R   N   T   A   K   E   K   K   K   V   R   D   T   D   M   L   M
ACCAACTGGCGGAACACGGCCAAGGAGAAGAAGAAAGTCCGAGACACGGACATGCTGATG                    1320

    A   Q   M   I   G   D   A   T   P   S   R   G   S   S   P   M   P   S   R   G
GCTCAGATGATCGGCGACGCAACACCGAGCCGAGGCTCGTCGCCGATGCCGAGCCGGGGC                    1380
```

Figure 2

Figure 2 (Continued)


```
S   S   P   V   H   L   L   H   K   G   M   G   R   S   D   D   P   Q   S   A
TCATCACCCGTGCACCTGCTTCACAAGGGCATGGGGCGGTCGGACGACCCCCAGAGCGCG          1440

P   T   S   P   R   T   Q   Q   E   A   R   D   M   Y   P   V   V   V   A   H
CCCACCTCGCCAAGGACCCAGCAGGAGGCTAGGGACATGTACCCGGTTGTGGTGGCGCAC          1500

P   V   H   R   L   N   P   N   D   R   R   R   S   A   S   S   S   A   L   E
CCGGTGCACAGACTAAATCCTAACGACAGGAGGAGGTCCGCCTCGTCGTCGGCCCTCGAA         1560

A   D   I   P   S   A   D   F   S   F   S   Q   G   *
GCCGACATCCCCAGTGCAGATTTTTCCTTCAGCCAGGGATGA          1602
```

Figure 4
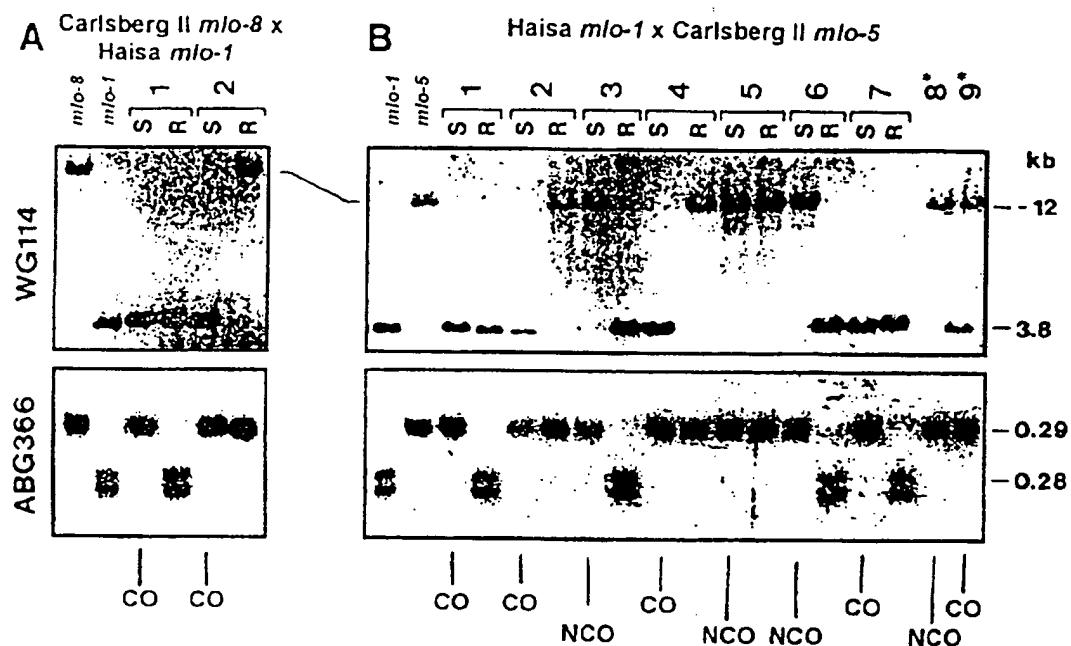
```
292 GCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCA 341
    || :||||| |||||| ||||||||||:|||||||| |||| ||| | ||
 80 GCANAGCTGATGCTGCTGGGCTTCATNTCCCTGCTTCTCACCGTGGCACA 129

342 GGACCCCATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCA 391
    || || |||| ||||:|| |||||| || || || ||| || ||
130 GGCGCC...CATCTCCAANATCTGCATCCCCAAGTCGGCTGCCAACATCT 176

392 TGTGGCCCTGCAAGCGCGGCACCGAGGGCCGC.AAGCCCAGCAAGTACGT 440
    ||| ||| ||||||| ||| : || | | ||| :| || | ||
177 TGTTGCCGTGCAAGGCAGGCCNAGATGCCATCGAAGAANAAGCAGCAAGT 226

441 TGACTACTGCCCGGAGGTGAGCAGCAGAGCCCGGACCAGCAGCTTCACGA 490
     | |: |:| || || || | | | | || | | ||| :|
227 GGTCNCCNGTCC.TTGGCCGGCGCCGGCGGCGGGGACTACTGCTCNAAAT 275

491 TGATGAAGAAATCAATACC...........GAACTTTTTCTTGTTTTCT 528
    |      || |  || :||        | ||     ||    : :  :
276 TCGATGTGAGAATAACNCCAGCTGCCGGCAAGCACAACCTCGATNCNATN 325

529 TCTGATTGTCGTCTTGGCTTGGCTTAATTGGTGTGTGTGTGTGTGTTTGC 578
    ||:||| |     |||||| | | | | | | ||| 
326 ACTNATT...........TAACTATAATTGATTTTTCTTGGGTTTTCTGC 364

579 AGGGCAAGGTGGCGCTCATGTCCACGGGCAGCTTGCACCAGCTGCACGTC 628
    ||||||||||||||| ||||| |   ||| ||||||||||||||| |
365 AGGGCAAGGTGGCGCTGATGTCGGCAAAGAGCATGCACCAGCTGCACATT 414

629 TTCATCTTCGTGCTCGCGGTCTTCCATGTCACCTACAGCGTCATCACCAT 678
    |||||||||||||||| || |||||||| |||||| || |||||||||
415 TTCATCTTCGTGCTCGCCGTGTTCCATGTTACCTACTGCATCATCACCAT 464

579 AGCTCTAAGCCGTCTCAAAGTGAGCCTTTGCTTCT.....TCTTCTTCTT 723
    | | || | || ||||||||||| || | |||| ||| | ||
465 GGGTTTAGGGCGCCTCAAAGTGAGTTTGTCGTTCTGTCCCTCATGCACAT 514

724 CTTTTACC........GCACGTCTGTCTGTCAGGCGTACCTACCTGTTCA 765
    |||| |        ||| :  |||| |||    | | |   |||
515 GTTTTCTCTAGTTCTAGCAANATTGTCAGTCCTTCAAATGGATTGTTTCG 564

766 TCAGGCTTGAGTAAAACTGTTCCATAATCTGC........TCCGGCATAA 807
    ||        || || | || |||| |||      |      ||||
565 ACA......AGAAACCCAATTTATTAATTTGCCAGTTAAATATATAATAA 608

808 TCCTCTCCTCCTG....CAGATGAGAACATGGAAGAAATGGGAGACAGAG 853
    |        || ||    |||||| | |||||||||| ||||| || ||
609 TTGATCTTTCTTGGTTTTAGATGAAGAAATGGAAGAAGTGGGAGTCACAG.658

854 ACCACCTCCTTGGAATACCAGTTCGCAAATGGTCAGGATCCCCCACTCTG 903
    |||| ||| |||||| || |||||||||| ||| |       | |
659 ACCAACTCATTGGAGTATCAGTTCGCAATCGGTAGTG.......AATTAA 701

904 CAATCTCCC...CTTCTTCGAAACCAAACC....TGATGATCCATTTAAA 946
    ||||||||   || |||     | ||||    |||| ||   || |||
702 GAATCTCCCTAACTATTTCATTTCAGAACCTTTATGATAATGTCTTGAAA 751

947 GACGCAGGCACGATCAGAGTGAGTGAACTGATGTATGTTCATTTTTTGTG 996
    || | |    ||||| ||| |   | ||||| |
752 GAGGAGGAGCAAATCAG.CTGAAAAATATGATCGA.............. 785

997 TCCTTTCAGATCCTGCACGGTTCCGGTTCACGCACCAGACGTCGTTCGTG 1046
    ||| | ||||||||| |||| ||| |||||||||| |||||||||||||||
786 TCCATGCAGATCCTTCACGATTCAGGTTCACGCATCAGACGTCGTTCGTG 835
```

FIGURE 5 cont'd

```
1047 AAGCGCCACCTGGG....CCTCTCCAGCACCCCTGGCATCAGATGGGTGGT 1093
     ||||| || |||||      |||| ||||||||||| |||||||| | ||
 836 AAGCGGCATCTGGGATCATTCTCAAGCACCCCTGGGCTCAGATGGATCGT 885

1094 GAGTTTTTTAGCTTCTTATCTGCCCCTCATCTGTGTGTAATGTT...... 1137
     |||||| | | ||| | ||       ||| ||| | | ||
 886 GAGTTATCAATCTCCGAAT.......ACATGCTTGTTTTTTATTCTTGCA 928

1138 ..TGGCGTA..........TGGAGTCAGGTGATTT.........ACCTT 1165
     |||| ||            | | || | |||          |  |
 929 ACTGGCCTAGCTGTTCCAATTCAATCCATATTTTTTGAAAAAAAAAATAT 978

1166 GCCTGTGATGTTTGTTGCCTTGTCAGGTGGCCTTCTTCAGGCAGTTCTTC 1215
     | || |||||| |||| |||| || ||||||||||||||||||
 979 TCATGCCGTGTTTG.....TTGTTAGGTAGCATTCTTCAGGCAGTTCTTT 1023

1216 AGGTCAGTCACCAAGGTGGACTACCTGACCTTGAGGGCAGGCTTCATCAA 1265
     |||| |||||||||||||||||||||||||||| || || ||||||||||
1024 GGGTCCGTCACCAAGGTGGACTACCTGACCATGCGGCAAGGCTTCATCAA 1073

1266 CGTACGTGC....CTCCCCTTCTAGCTCCGCCATTGCTGCCGCGATGTAG 1311
     ||| | |    |  |||    | | ||||| || || ||
1074 TGTATATACTAATCAAACCTGACCAATTCAACATTGATGATGC.AAACAG 1122

1312 CAGCAAAGCTTCT.......CAAGTTATCCTTCTGACGCTAAAGTTCCCA 1354
     || || || |       | |||| | | ||| | ||| |
1123 AAGACCAGGTTTTTTTTTTCCGAGTTGTGCAT.TGAAGTTAATG...... 1165

1355 TGTTTTTTCCTCAAATTATTCTGCGCAGGCG.CATTTGTCGCAAAACAGC 1403
     ||||| | || || | |  |||||||| |||||||||||| || |||
1166 .GTTTAGCTTC...TTCTCTTTTGCAGGCGCCATTTGTCGCAGAATAGC 1211

1404 AAGTTCGACTTCCACAAGTACATCAAGAGGTCGATGGAGGACGACTTCAA 1453
     ||||||||||||||||||| ||||||||||| ||||||||||||||||||
1212 AAGTTCGACTTCCACAAATACATCAAGAGGTCTTTGGAGGACGACTTCAA.1261

1454 GGTCGTCGTCGGCATCAGGTACGTTCCATTCCTTCCTCTGCACCAGACCA 1503
     || ||||| ||||||||| ||     |||| || |
1262 AGTTGTCGTTGGCATCAGGTCCG......TCCTCGCTTT........... 1294

1504 CACCCCATGGATAGATTTTAACAATTGCTGTCAGGTTCCACATGATAACA 1553
     || ||| || |    || | | ||| ||||| |
1295 ........ATTAATTATAGGA....CTCTTATATTCAACATTTTTTT 1330

1554 ATATACTATGA.ACTTGGTCTTTGCTCCTTGTCCTTG.....CACGATCA 1597
     ||| | | | || |||| |    |||| ||      ||||
1331 ATAAAGAAACATATTTAGTCT...CCAGTTGTGTATGTGTATGTGGATCT 1377

1598 TGACACATTTGGCCTGTTTTCGCAGCCTCCCGCTGTGGGGTGTGGCGATC 1647
     ||||||||||||| ||| ||| ||||||||| ||||||| || | |||
1378 TGACACATTTGG.CTGGTTTTGCAGCCTCCCTCTGTGGTTCGTCGGAATC 1426

1648 CTCACCCTCTTCCTTGACATCAATGGTATGGACCTTCTCCTCTCCGGTTT 1697
     ||  ||||||||| || ||| | |||    |||| ||||       |||
1427 CTTGTACTCTTCCTCGATATCCACGGTA..ATCCTTGTCCT.....ATTT 1469

1698 CTCTATTGCTTTGCAGCTAAATAAAACACTTGCAATTCGTCTCGTGATCA 1747
     | | || |||   || || |    | ||||| |||| ||| | |||||
1470 CATTCTTTTTTTTACTCTCAAAACCTTGTTCTGAATTGGTCTTATAATCA 1519

1748 CCGCTCATTTTTCAACCATTTCTTTTTCTACTCATAGGGGTTGGCACGCT 1797
     ||  |||||| | | || ||| | |    |||| ||||||| ||
1520 CCATCGATTTTTTTTCAACTT.TTTCCCCGCGTGTAGGTCTTGGCACACT 1568

1798 CATCTGGATTTCTTTCATCCCTCTCGTGGTAAGTGC.AGATTTCTCC.AT 1845
     ||  ||||| ||||| |  ||||||| |||||| || | ||||| || |
1569 TATTTGGATCTCTTTTGTTCCTCTCATCGTAAGAGCGAAATTTCCCCTGT 1618
```

FIGURE 5 cont'd

```
1846 CGAAAGCAACAGCAAACCCAATT...............TGATCGCAAT 1878
     I IIII IIIII  III  IIII           I II   III
1619 CCAAAGAAACAGTTAACATAATTAATTATGCTTTAATTTATCATGAAAAT 1668

1879 GGAAACCCACACCTAATATTAACTCAAAATGTCAATTGTCGGTGCGTCTT 1928
     I I   II  III       IIII   III  III   III I
1669 TAATATGATCATATAACTAATGAACAAACATTCA..TGTGAATGCCACCG 1716

1929 CCTCAACAGATCCTCTTGTGTGTTGGAACCAAGCTGGAGATGATCATCAT 1978
     IIIIII IIIIIII  IIIII IIIIIIIII IIIIII I IIIII
1717 TTGTCTCAGATCGTCTTGTTAGTTGGGACCAAGCTAGAGATGGTGATCAT 1766

1979 GGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCGTCATCAAGGGGGCCC 2028
     IIIIIIIIII IIIII IIIIII IIII I  II III IIII II I
1767 GGAGATGGCCCAAGAGATACAGGACAGGGCCACTGTGATCCAGGGAGCAC 1816

2029 CCGTGGTCGAGCCCAGCAACAAGTTCTTCTGGTTCCACCGCCCCGACTGG 2078
     I  IIII II II IIIIIIIIII IIIIIIIIII IIIIIII IIIIII
1817 CTATGGTTGAACCAAGCAACAAGTACTTCTGGTTCAACCGCCCTGACTGG 1866

2079 GTCCTCTTCTTCATACACCTGACGTTGTT................. 2107
     III I IIIIIIIIIIIIIIIII I II
1867 GTCTTGTTCTTCATACACCTGACACTCTTCCCATGTACATGTTTAAAACC 1916

                          .
                          .
                          .

2108 ...................CCAGAACGC.GTTTCAGATGGCGCATTTG 2136
                       IIIIIIIII IIIIIIIIIIIIIIIIIII I
2017 GACGGACGGATCGATCATCACCAGAACGCATTTTCAGATGGCGCATTTCG 2066

2137 TGTGGACAGTG....GTACGCCAC.........CGATGAACTTGTCAGTT 2173
     I IIIIII II   III II II        I II  II IIIIII
2067 TATGGACTATGGTGTGTATGCTACTTGCTTAGTTGTTGCCATTATCAGTT 2116

2174 .......AACATGGGTGTCA...AGGCACCGAGTGCCGCTGATGA..... 2208
            II I IIII I   I IIII II    II IIII
2117 CTTAAGCAAATTAAGTGTGATGCATGCACTGA......CTAATGAGACAA 2160

2209 ........ACTGCTCTGACGGAGATTTACTTGTGTTGT.......AGGCC 2243
             II III    I   III I IIII IIII      IIII
2161 AAAATGACACAGCTTGTTCATCGATCTGGTTGTTTTGTGTGTGACAGGCA 2210

2244 ACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGGCTGAGCATCAT 2293
     II II II IIIIIIIIIIII III    I II IIIIIIIIII I
2211 ACACCTGGTCTGAAGAAATGCTTCCATGAAAATATTTGGCTGAGCATCGT 2260

2294 GAAGGTGGTGGTGGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCT 2343
     I I II I IIIIII I IIII III I II IIIIIIII II IIII
2261 GGAAGTCATTGTGGGGATCTCTCTTCAGGTGCTATGCAGCTACATCACCT 2310

2344 TCCCCCTCTACGCGCTCGTCACACAGGTAATAAAACCGTCCAGGAA 2389
     IIII IIIIIIIIIIIIIIIIIIIIIIIII I  II II I II  II
2311 TCCCGCTCTACGCGCTCGTCACACAGGTGAACAAGCCATTCACAAA 2356
```

FIGURE 6

```
295 GAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCAGGA 344
    ||||||:|||||||||||||||||||||||||||||||||||||||||||
  1 GAGCTCNTGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCAGGA 50

345 CCCCATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCATGT 394
    ||  | || ||| ||| |||||||||| |||| |||| |||:  :|||
 51 TCC...CGTCTCCAGGATCTGCATCTCCAAGGAGGCCGGCGANAANATGC 97

395 GGCCCTGCAAG.....CGCGGCACCGAGGGCCGCAAGCCCA........ 430
    || ||||||    : |||| ||| || |||| ||
 98 TCCCGTGCAAGCCTTACNACGGCGCCGGCGGTGGCAAAGGCAATGACAAT 147

431 ........................GCAAGTACGTTGACTACTGCCCGGA 455
                            || |: || : | || ||| |
148 CACCGGAGGCTTCTCTGGCTCCAAGGCGANAGCGANACCCACCGCCGGTT 197

456 GGTGAGCAGCAGAGCCCGGACCAG.................. 479
    || || ||  | |||||: |
198 CCTG.GCTGCCCCGGCCGGANTGGACGTCTGCGCCAAACAGGTGAGCACC 246

480 CAGCTTCACGATGATGAAGAAA.TCAATACCGAACTTTTTCTTGTTTTCT 528
    |:| ||:| | |  | ||| | : || || | | || ||
247 TANCGTCNCCACAAACCACAAACTANCTAATGAGCATGGACCTGAATTTC 296

529 TCTGATTGTCGTCTTGGCTTGGCTTAATTGGTGTGTGTGTGTGTGTTTGC 578
    |   || | | |||||||| || |||||||                | |||
297 TTCTCTTCTTGGCTTGGCTTGACTAAATTGGT.............TGTGC 333

579 AGGGCAAGGTGGCGCTCATGTCCACGGGCAGCTTGCACCAGCTGCACGTC 628
    | |||||||||||||| |||||::|||| |:| ||||||| |||||| |
334 ACGGCAAGGTGGCGCTGATGTCNNCGGGAANCATGCACCAACTGCACATA 383

629 TTCATCTTCGTGCTCGCGGTCTTCCATGTCACCTACAGCGTCATCACCAT 678
    |||||||||||||||||| |||||| ||| ||||||||| |||||||
384 TTCATCTTCGTGCTCGCCGTCTTCCACGTCTTGTACAGCGTCGTCACCAT 433

679 AGCTCTAAGCCGTCTCAAAGTGAGCCTTTGCTTCTTCTTCTTCTTTT 728
    | ||||||||||||||||||||||| |     ||
434 GACCCTAAGCCGTCTCAAAGTGAGCATCATACTC.............. 467

729 ACCGCACGTCTGTCTGTCAGGCGTACCTACCTGTTCATCAGGCTTGAGTA 778
       |||| |||||   .| | |   ||| |  |       |
468 ......GAGCTGTTTGTCAATAATCCTT...GGTTTCCAATCCAATTCCA 508

779 AAACTGTTCCATAATCTGCTCCGGCATAATCCTCTCCTCCTGCAGATGAG 828
    || |||  |  | ||||||||          | |||||||||||
509 AAGCTGGCACTGATCCTGCTCCGG...........CTTCCTGCAGATGAA 547

829 AACATGGAAGAAATGGGAGACAGAGACCACCTCCTTGGAATACCAGTTCG 878
    ||||||||| ||||||  |  ||||||  ||||   |||| || |||||||
548 GCAATGGAAGAAGTGGGAGTCGGAGACCGCCTCGCTGGAGTATCAGTTCG 597

879 CAAATGGTCAGGATCCCCCACTCTGCAATCTGCCCTTCTTCGAAACCAAA 928
    | ||||||||||                ||| |  | ||| |    |||
598 CGAATGGTCAG..................CTTCAACTTTTCTTACTGAAA 629

929 CCTGATGATCCATTT...AAAGACGCAGGCACGATCA.....GAGTGAGT 970
    || ||||  ||||| | | |||||| |||||||||     | | ||
630 CCGGATG...CATTTACAACAAACGCACGCACGATCAATCATCACAGTGT 676

971 GAACTGAT.GTATGTTCATTTTTTGTGTCCT.TTCAGATCC..TGCACGG 1016
    || | |||     || | ||  |||| |||||||| |   |||
677 GAGCCGATACGTTGAACCCGATTGAAATCCTCCGCAGATCCCATCGCCGG 726
```

```
1017 TTCCGGTTCACGCACCAGACGTCGTT.CGTGAAGCGCCACCTGGGCCTCT 1065
     I IIIIIIIIIIIIIIIIIIIIIII IIII  IIII III IIIIIIIIIIIII
 727 TGCCGGTTCACGCACCAGACGACGTTGGGTGAGGCGGCACCTGGGCCTCT 776

1066 CCAGCACCCCTGGCATCAGATGGGTGGTGAGTTTTTTAGCTTCTTATCTG 1115
     IIIIIIIIIII III IIIIIIIIII
 777 CCAGCACCCCCGGCGTCAGATGGGT.................... 801


1166 GCCTGTGATGTTTGTTGCCTTGTCAGGTGGCCTTCTTCAGGCAGTTCTTC 1215
                       IIIIIIIIIIIIIIIIIIIIIIIIIIIIII
 802 ...................GGTGGCCTTCTTCAGGCAGTTCTTC 826

1216 AGGTCAGTCACCAAGGTGGACTACCTGACCTTGAGGGCAGGCTTCATCAA 1265
     I IIII II IIIIIIIIIIIIIIIIIIIIIIIII II  IIIIIIIIIII
 827 ACGTCGGTGACCAAGGTGGACTACCTGACCTTGCGGCAGGGCTTCATCAA 876

1266 CGTACGTGCCTCCCCTTCTAGCTCCGCCATTGCTGCCGCGATGTAGCAGC 1315
     I
 877 C.................................................... 877



1366 CAAATTATTCTGCGCAGGCGCATTTGTCGCAAAACAGCAAGTTCGACTTC 1415
                      IIIIII I IIIII  II II IIIIIIIIII
 878 ................GCGCATCTCTCGCAGGGCAACAGGTTCGACTTC 910

1416 CACAAGTACATCAAGAGGTCGATGGAGGACGACTTCAAGGTCGTCGTCGG 1465
     IIIIIIIIIIIIIIIIIIIIIIII IIIIIIIIIIIIIIIIII IIIIIIIII I
 911 CACAAGTACATCAAGAGGTCGTTGGAGGACGACTTCAAAGTCGTCGTCCG 960

1466 CATCAGGTACGTTCCATTCCTTCCTCTGCAC........CACACCACAC 1506
     IIIIIIIIIII IIIIIIIII IIIIIIII        II I  III
 961 CATCAGGTACGCGCCATTCCTTTCTCTGCACAAATTAATACATCCACCAC 1010

1507 CCCATGGATAGATTTTAACAATTGCTGTCAGGTTCCACATGATAACAATA 1556
     I III : IIIII    I I              II :  I:I II
1011 CACATANGTAGATAGATAGA................TCGATANATANATTA 1045

1557 TACTATGAACTTGGTCTTTGCTCCTTGTCCTTGCACGATCATGACACATT 1606
     III I I  I I I II I   I IIII IIIIIIII
1046 TAC.AAGTGCCGGTACGTACGTACGTCTCAT...ATGATCTTGACACATC 1091

1607 TGGCCTGTTTTCGCAGCCTCCCGCTGTGGGGTGTGGCGATCCTCACCCTC 1656
     II III II  IIII  III  III III    IIIIII IIIIII IIII
1092 TGTCCTCTTGCCGCAATCTCAAGCTCTGGTTCGTGGCGGTCCTCATCCTC 1141

1657 TTCCTTGACATCAATGGTATGGACCTTCTCC.TCTCCGGTTTCTCTATTG 1705
     IIIIIIII II I IIII   IIIII III I II II  I   I  I
1142 TTCCTTGATTTCGACGGTAGCCGCCTTGTCCATGCCCTGCTCGCCCTCTC 1191

1706 CTTTGCAGCTAAATAAAACACTTGCAATTCGTCTCGTGATCACCGCTCAT 1755
     II  II  II   I II   II II I III III          II
1192 CTCCGCTTCTCTCCATAATTTGTG.AACTTGTCCCGT...........AT 1229

1756 TTTTCAACCATTTCTTTTTCTACTCATAGGG.TTGGCACGCTCATCTGG 1804
     I I II    I I  III III IIIII I IIIII II  IIIII
1230 ATAACCACACCACCGTCGTCTTCTCGCAGGGGATCGGCACTCTTCTCTGG 1279

1805 ATTTCTTTCATCCCTCTCGTGGTAAGTGCAGATTTCTCCATCGAAAGCAA 1854
     II II  I  I IIIIIIIIIIIIIIIII II     I II  II   I
1280 ATGTCCGTGGTTCCTCTCGTGGTAAGTCCA.......CAATTTGAATAGA 1322

1855 CAGCAAACCCAATTTGATCGCAATGGAAACCCACACCTAATATTAACTCA 1904
     II I   IIIIII   I I I I  III II III       II
1323 CAACCTGTCCAATTGTGATGTACAGTACCTCCAAACTTAA.......TTA 1365
```

```
1905 AAATGTCAATTGTCGGTGCGTCTTCC.....TCAACAGATCCTCTTGTGT 1949
     I IIIIIII III  I I  IIIII I      II IIIIIIIIIIIII
1366 ACATGTCATTTGCTGAT..GTCTTGCGTGTAACATTAGATCCTCTTGTGG 1413

1950 GTTGGAACCAAGCTGGAGATGATCATCATGGAGATGGCCCTGGAGATCCA 1999
     IIIIII IIIIIIIIIIIIIIII I IIIIIIIIIIIIIIII III:IIIII
1414 GTTGGGACCAAGCTGGAGATGGTGATCATGGAGATGGCCCAGGANATCCA 1463

2000 GGACCGGGCGAGCGTCATCAAGGGGGCCCCCGTGGTCGAGCCCAGCAACA 2049
     IIIIIII IIIIIIII IIIIIII II IIII IIIIIIIIIIIIIIII
1464 TGACCGGGAGAGCGTCGTCAAGGGTGCTCCCGCCGTCGAGCCCAGCAACA 1513

2050 AGTTCTTCTGGTTCCACCGCCCCGACTGGGTCCTCTTCTTCATACACCTG 2099
     III IIIIIIIIII IIII II IIIIIIIIIIIIIIII IIII IIIII
1514 AGTACTTCTGGTTCAACCGGCCTGACTGGGTCCTCTTCCTCATGCACCTC 1563

2100 ACGTTGTTCCAGAACGCGTTTCAGATGGCGCATTTTGTGTGGACAGTGGT 2149
     II I IIIIIIIIIIIIIIIIII IIIII IIIIIIIIIIIIII
1564 ACACTCTTCCAGAACGCGTTTCAGATGGCTCATTTCGTGTGGACAGTGGT 1613

2150 ACGCCACCGATGAACTTGTCAGTTAACATGGG................. 2181
     I  I:  I I IIIIIIII II II I :I
1614 A...CNTACAAGTACTTGTCACTTCACTTANGCTAACTCCAACAAACGAA 1660


2182 ........TGTCAAGGCACCGAGTGCCGCTGATGAACTGCTCTGACGGAG 2223
             IIII II II I I   I  I I  II I
1711 GACACAAAACTCAATCCAACGCGCGGTAGCAAACGAACGTTTTTCCGTAC 1760

2224 ATTTAC.........................................TTG 2232
     III I                                        III
1761 GTTTTCGTCCGCTTTCGCCCCATCCCAGCCCAAATTCGTTGACGTTGTTG 1810

2233 TGTTGTAGGCCACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGG 2282
     I I IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII I I II I
1811 CATCGCAGGCCACGCCCGGCTTGAAGAAATGCTACCACGAGAAAATGGCA 1860

2283 CTGAGCATCATGAAGGTGGTGGTGGGGCTAGCTCTCCAGTTCCTCTGCAG 2332
     IIIIIIII  IIIII III IIIII IIII  IIII II I IIIII
1861 ATGAGCATCGCCAAGGTCGTGCTGGGGGTAGCCGCCCAGATCTTGTGCAG 1910

2333 CTATATGACCTTCCCCCTCTACGCGCTCGTCACACAGGTAATAAAACCGT 2382
     :II II IIIIIIIII II:IIIIIIIIIIIIII
1911 NTACATCACCTTCCCGCTNTACGCGCTCGTCAC................. 1943


2433 AATCATCTGTGTGTGCTGGCTTTGTATGCAGATGGGATCAAACATGAAGA 2482
                             IIIIIIIIII III IIIIIIIII
1944 ............................GCAGATGGGCTCACACATGAAGA 1966

2483 GGTCCATCTTCGACGAGCAGACGTCCAAGGC.GCTCACCAACTGGCGGAA 2531
     I   II:IIIIIIIIIIIIIII IIIIIII III IIIIIIIIIII II
1967 GAAGCANCTTCGACGAGCAGACGGCCAAGGCGGCTGACCAACTGGCGAAA 2016

2532 CACGGCCAAGGAGAAGAAGAAAGTCCGAGACACGGACATGCTGATGGCTC 2581
     I IIIIIIIIIIIIIIIIIIIII I IIIIIII III IIIIIIIIIIII I
2017 GATGGCCAAGGAGAAGAAGAAGGCCCGAGACGCGGCCATGCTGATGGCGC 2066

2582 AGATGATCGGCGACGCAACACCGAGCCGAGGCTCGTCGCCGATGCCGAGC 2631.
     IIIII IIIIII III II IIIIII IIII:IIIIIII
2067 AGATGGGCGGCGGCGCGACGCCGAGCGTCGGCTNGTCGCCG........ 2107
```

SUBSTITUTE SHEET (RULE 26)

FIGURE 6 cont'd

```
2632 CGGGGCTCATCACCCGTGCACCTGCTTCACAAGGGCATGGGGCGGTCGGA 2681
          ||||||||||| ||||||| |  || ||||||  ||
2108 ..............GTGCACCTGCTCCACAAGGCCGGGGCGCGGTCCGA 2142


2682 CGACCCCCAGAGCGCGCCCACCTCGCCAAGGACCCAGCAGGAGGCTAGGG 2731
     ||||||||||||| |||  | || || ||| || || ||||  |    | |
2143 CGACCCCCAGAGCGTGCCGGCGTCCCCGAGGGCCGAGAAGGAAGGCGGCG 2192


2732 ACATGTACCCGGTTGTGGTGGCGCACCCGGTGCACAGACTAAATCCTAAC 2781
      |              |||   |||  ||||  ||  ||   ||   ||||
2193 GC...............GTGCAGCATCCGGCGCGCAAGGTACCTCCTTGT 2227


2782 GACAGGAGGAGGTCCGCCTCGTCGTCGGCCCTCGAAGCCGACATCCCCAG 2831
     |||  || |||||||| ||||||||| |||| ||||| ||  |||||||| |
2228 GACGGGTGGAGGTCGGCCTCGTCGCCGGCGCTCGACGCTCACATCCCCGG 2277


2832 TGCAGATTTTTCCTTCAGC.........CAGGGATGAGACAAGTTTCTG 2871
     |||||||||| |||||||            |  ||| ||||||||| ||
2278 TGCAGATTTTGGCTTCAGCACGCAACGTTGACCGATCAGACAAGTTCCTT 2327


2872 TATT 2875
     | ||
2328 TTTT 2331
```

Figure 7   13/28

```
                              ι ι        .            ∎
                         GGCTGCTCCGCCAGCAAACCAGACACACAGCAGCGTACCTGCGT

                    ACGTAGCGTGCGCTTTCTTTTTTTTTCCTTTCGCCTCTCTTGCTTGCTCCGGCCGGCCACG

                    TCGATAGCCGGCCACGGCCAGGCACCTCGCGGTTGCGTCGCGTGCATCTGCGTGTGCGTA

                    CCTGGTAGAGGCGGCCGTCTGCTTGCTCCGGGCAAGGAAGGAGCGCTTGCGGCGGTCGACCG

              M  S  D  K  K  G  V  P  A  R  E  L  P  E  T  P  S [W   A  V  V]    20
         ATGTCGGACAAAAAAGGGGTGCCGGCGCGGGAGCTGCCGGAGACGCCGTCGTGGGCGGTG         60

        [A  V  V  F  A  A  M  V  L  V  S  V  L  M] E  H  G  L  H  K           40
         GCGGTGGTCTTCGCCGCCATGGTGCTCGTGTCCGTCCTCATGGAACACGGCCTCCACAAG        120

              L  G  H  W  F  Q  H  R  H  K  K  A  L  W  E  A  L  E  K  M       60
         CTCGGCCATTGGTTCCAGCACCGGCACAAGAAGGCCCTGTGGGAGGCGCTGGAGAAGATG        180

              K  A  E [L  M  L  V  G  F  I  S  L  L  L  I  V  R  T  Q  D  A E]  80
         AAGGCGGAGCTCATGCTGGTGGGCTTCATATCCCTGCTCCTCATCGTCACGCAGGACCCC        240

        [I  R  I  A  K  I  C  I  S] E  D  A  A  D  V  H  W  P  C  K  R        100
         ATCATCGCCAAGATATGCATCTCCGAGGATGCCGCCGACGTCATGTGGCCCTGCAAGCGC        300

              G  T  E  G  R  K  P  S  K  Y  V  D  Y  C  P  E  G  K  V  A      120
         GGCACCGAGGGCCGCAAGCCCAGCAAGTACGTTGACTACTGCCCGGAGGGCAAGGTGGCG        360

              L  M  S  T  G  S  L  H  Q  L  H [V  F  I  F  V  L  A  V  V F]    140
         CTCATGTCCACGGGCAGCTTGCACCAGCTGCACGTCTTGATCTTCGTGCTCGCGGTCTTC        420

        [H  V  T  Y  S  V  I  T  I  A  L] S  R  L  K  M  R  T  W  K          160
         CATGTCACCTACAGCGTCATCACCATAGCTCTAAGCCGTCTCAAAATGAGAACATGGAAG        480

              K  W  E  T  E  T  T  S  L  E  Y  Q  F  A  N  D  P  A  R  F     180
         AAATGGGAGACAGAGACCACCTCCTTGGAATACCAGTTCGCAAATGATCCTGCACGGTTC        540

              R  F  T  H  Q  T  S  F  V  K  R  H  L  G  L  S  S  T  P  G     200
         CGGTTCACGCACCAGACGTCGTTCGTGAAGCGCCACCTGGGCCTCTCCAGCACCCCTGGC        600

              I  R  W  V  V  A  F  F  R  Q  F  F  R  S  V  T  K  V  D  Y     220
         ATCAGATGGGTGGTGGCCTTCTTCAGGCAGTTCTTCAGGTCAGTCACCAAGGTGGACTAC        660

              L  T  L  R  A  G  F  I  N  A  H  L  S  Q  N  S  K  F  D  F     240
         CTGACCTTGAGGGCAGGCTTCATCAACGCGCATTTGTCGCAAAACAGCAAGTTCGACTTC        720

              H  K  Y  I  K  R  S  M  E  D  D  F  K [V  V  V  G  I  S  L]     260
         CACAAGTACATCAAGAGGTCGATGGAGGACGACTTCAAGGTCGTCGTCGGCATCAGCCTC        780

        [P  L  W  G  V  A  I  L  T  L  F  L] D  I  N  G  V  G [T  V  L]      280
         CCGCTGTGGGGTGTGGCGATCCTCACCCTCTTCCTTGACATCAATGGGGTTGGCACGCTC        840

        [I  W  I  S  F  I  P  L  V  I  L  L  C  V  G] T  K  L  E  M         300
         ATCTGGATTTCTTTCATCCCTCTCGTGATCCTCTTGTGTGTTGGAACCAAGCTGGAGATG        900

              I  I  M  E  M  A  L  E  I  Q  D  R  A  S  V  I  K  G  A  P     320
         ATCATCATGGAGATGGCCCTGGAGATCCAGGACCGGGCGAGCGTCATCAAGGGGGCCCCC        960

              V  V  E  P  S  N  K  F  F  W  F  H  R  P  D  W  V  L  F  F     340
         GTGGTCGAGCCCAGCAACAAGTTCTTCTGGTTCCACCGCCCCGACTGGGTCCTCTTCTTC       1020

              I  H  L  T  L  F  Q  N  A  F  Q  H  A  H  F  V  W  T  V  A     360
         ATACACCTGACGTTGTTCCAGAACGCGTTTCAGATGGCGCATTTTGTGTGGACAGTGGCC       1080

              T  P  G  L  K  K  C  Y  H  T  Q  I  G  L  S  I  M  K [V  V]     380
         ACGCCCGGCTTGAAGAAATGCTACCACACGCAGATCGGGCTGAGCATCATGAAGGTGGTG       1140

        [V  G  L  A  D  Q  F  L  C  S  Y  H  T  F  P  L  T  A  L  V]         400
         GTGGGGCTAGCTCTCCAGTTCCTCTGCAGCTATATGACCTTCCCCCTCTACGCGCTCGTC       1200

        [T] Q  M  G  S  N  M  K  R  S  I  F  D  E  Q  T  S  K  A  L         420
         ACACAGATGGGATCAAACATGAAGAGGTCCATCTTCGACGAGCAGACGTCCAAGGCGCTC       1260

              T  N  W  R  N  T  A  K  E [K  K  K  V  R] D  T  D  H  L  M     440
         ACCAACTGGCGGAACACGGCCAAGGAGAAGAAGAAAGTCCGAGACACGGACATGCTGATG       1320

              A  Q  M  I  G  D  A  T  P  S  R  G  S  S  P  M  P  S  R  G     460
         GCTCAGATGATCGGCGACGCAACACCGAGCCGAGGCTCGTCGCCCGATGCCGAGCCGGGGC       1380

              S  S  P  V  D  L  L  K  G  M  G  R  S  D  D  P  Q  S  A       480
         TCATCACCCGTGCACCTGCTTCACAAGGGCATGGGGCGGTCGGACGACCCCCAGAGCGCG       1440

              P  T  S  P  R  T  Q  Q  E  A  R  D  M  Y  P  V  V  V  A  H    500
         CCCACCTCGCCCAAGGACCCAGCAGGAGGCTAGGGACATGTACCCGGTTGTGGTGGCGCAC       1500

              P  V  H  R  L  N  P  N  D  R  R  R  S  A  S  S  S  A  L  E    520
         CCGGTGCACAGACTAAATCCTAACGCACAGGAGGAGGTCCGCCTCGTCGTCGCCCTCGAA       1560

              A  D  I  P  S  A  D  F  S  F  S  Q  G  *
         GCCGACATCCCCAGTGCAGATTTTTTCCTTCAGCCAGGGATGAGACAAGTTTCTGTATTCA

                    TGTTAGTCCCAATCTATAGCCAACATAGGATGTGATGATTCGTACAATAAGAAATACAAT
                                                               ι ι
                    TTTTTACTGAGTC
```

Figure 8

```
   1   GAATTCAATT  AAGGACAACA  ACGGATGATA  GGCTTAAGCT  AGAGAGGATT
  51   CATATCGATT  AATTAACTGT  ACTTAAGTTG  AGGTAAAACT  CTATCGATTG
 101   CTTTGGACAC  CGGCTCTCCC  ATGATCTGCC  AAGTTGAGCC  GGCCTACCTA
 151   ATTTTCTTCG  AAAGCACACA  ACAAACGAAG  GTAACCACTA  ATCTAGACAC
 201   CACGCCTAAG  TTATCAATTA  CTACTCTAGT  CTCGCGTAGA  AACTTCATTC
 251   TTTATGGAGA  GTGCTAGTAC  TAGAGTACTT  AATATAATAG  TAAGCGACAA
 301   ACCCACGACG  ATGAGAATGT  ACCTCACTTA  CGTAGTCAAT  TAAGTCGAAA
 351   AGGAAATCTT  GAACACTTAC  TTTATTAAAG  AAGTATTCCC  CGAGGTACAG
 401   GAGAGGAGAG  CACGCCAATA  ACTCCAGCAC  TCCTCCGAAA  CCTTTCTCAC
 451   TCTCTACCCT  TTTTCTCCAC  ACAACTAAAA  TGATGTCTAA  TGTATGAAAG
 501   TGAGTTGTAC  TCTATTTTGT  TGTGTGTTTG  GAAGTGAAAT  TAGCTCATCC
 551   TTTTATAGCA  ACTTAATGGT  CGGTTGTAGG  TTGGTAATTA  AGTCGGTAAA
 601   CACTCACAAC  CACCATCGTC  AACCAATAGG  AGATCGCCAC  ATGATCGAAA
 651   GCTGACAGTT  AGGGGTGCCA  ACCCTGTTTT  GTCCGAACCA  AGCAAACAAC
 701   CTCTATCTAG  GACCTCTCTT  CTATGTCTGA  CAAGTCGGCC  CATATGGCGG
 751   TGCACTATGG  ATTAAGTCAA  TTTCAGTCGT  TTTGGACTGT  CATGTGGGCC
 801   CTTCCAATCC  TTGTGCTCCC  ATATGATTGG  TCGAAAGTAC  ATTTAATTCC
 851   TGGGTGAGTG  CTAGAACTAA  TATGATAGAT  GTGCTCCGGC  TCCTGGGAAA
 901   GAGGCCACTT  GACATACTTG  GGGTAGTGCC  CCAAGGGTAT  TCCCTATCGC
 951   TTTTTCATAA  TTTTCTCTCT  CCAAAATCGG  ACGGAAACAA  TAAAAAGAG
1001   AGGCGATGTT  CATCGGCAAA  TATCTATTTT  TTTGATAGTG  TCTTCCCTTA
1051   AAACTTGATT  TTTGCGAAGA  CTTCCGGCTA  AAACCATGAA  ATCAGAGTTC
1101   CTTGTAACAA  ATTTAATTTG  CCTAAATACA  AAAAGATCG  AATGGAGATA
1151   GCATTAAACT  TGCTCCATAC  GAATCATATT  AGTTGGACCG  TAACTCATAG
1201   AAAAAGTTGC  AAGTTGGTTG  ACCTATCAAC  CCTCTTATGT  TGACCGTAAA
1251   CCTGTTATGC  ATTAAGGATT  AAGTACCGGC  AGATCGTCAC  TACTCACGAA
1301   TGCACAAATT  TCCGGTAACG  TAGGATGGGA  TGAGTTGGTC  AGAAACGGGT
1351   CACCACGTCG  CCCAACCTGC  CGCGATCGAG  CCATTGGCCG  GCGATGCACG
1401   CGCTTTGACA  CAGCCGCCCG  CCGCCCCCG  GCCCGCCCCC  GTTTTTAATA
1451   AAAACCGGCC  GCCCCCTGTC  AAAGGTCTCA  AAGTGTCAAG  TGCATCAGAG
1501   CTAAGCTAGC  GGTCACCCAG  TCAGCTCACC  CCGAGACGCA  CCAGGGATC
1551   TATCGGATCA  TGGCAGGTGG  GAGATCGGGA  TCGCGGGAGT  TGCCGGAGAC
1601   GCCGACGTGG  GCGGTGGCCG  TCGTCTGCGC  CGTCCTCGTG  CTCGTCTCCG
1651   CCGCCATGGA  GCACGGCCTC  CACAACCTCA  GCCATGTACG  CGCGCGCGCA
1701   CGCGGTGTGC  TCATCTCTCG  AGTTAATTTG  GTTGTTGTTG  TTGTTGTGTT
1751   CTTGTGACAT  CTCAATTAAC  ATCCGATCGT  GGTCGATCGA  TCGCCCTGTG
1801   GTGGCGATAC  TGCTTGCATT  GCAGTGCTTC  CGTAGGCGGC  AGAAGAAGGC
1851   CATCGGCGAC  GCCCTCGACA  AGATCAAAGC  AGGTCACCCT  CAGCCTCAGC
```

FIGURE 8 cont'd

```
1901  TCACCCTCAG CCTCCATCTC TAAATATTTG ACGCCGTTGA CTTTTTTAAA
1951  TATGTTTGAC CATTCGTCTT ATTTAAAAAA TTTAAGTAAT TATTAATTCT
2001  TTTTCTACCA TTTGATTCAT TGCTAAATAT ACTATTATGT ATACATATAG
2051  TTTTACATAT TTCACTAAAG TTTTTAAATA AGACGAATGG TCAAACATGT
2101  TTAAAAAAGT CAACGGCGTC AAACATTTAG GAAGAAGAGA ATATTATATT
2151  GCTGCTCCCC TCTAGCCACT TTGCTGCCTC CCTCGTCATT TTTTCAAGTA
2201  TTTTACGCAA GACTGGTCCT CCAAATCAAA CGTCACAAAT AAGCCATTTA
2251  TAGTTTCCTT TCGCTTTTTA AGGGGGACTA CTTGTATTTA ATCATGGAGG
2301  AAACTACCAG TCGGATGTCC GATTACTTAA AAAAAAATTC GGGGGACTAA
2351  TTTTTTTGGC TGATCATCGG TGAAATATTA GGTTATATAT GTTGAAAAAA
2401  AATCAGCCAC AAACAATGAA ATATTTTGTG AAACACATAT TAGACACGTT
2451  GAAACGTATC ATTGTTACGT ATAAAACATC GAATGTTAAC AGATTAAAAC
2501  ATATGTTTTT TTTTAATCAG AATATAATCA TGCGATATAT TATTGTAAAG
2551  ATATAATTAC AACGAATACA ACAGTGCGAT CGGATTATAT ATATATTAGT
2601  AGTTTAAGAG AAAAATCATT TTGAAGATTA CTAGATACAT ACACGTATAG
2651  ATGGATGAAG TGGAGAGAGA TTAGAGATAA GTAGTTATAT GAATTTTGTG
2701  AAACACACTT AAGACATATG TTCAAACATA CTGCTATTAT GTATGAAATA
2751  TTGAGTTTTA ACGGTTTAAA ACACATATTC TTTTAATTAG AATGTAATAA
2801  TGTGATATCT TGTTGTAAAA TTTAATTACA TCTAATATAA CGGTGTGATT
2851  AGATTGTATG TTGGATAACA TGCCCATCGG TTGGCTTATT TAGGGAATAA
2901  GCCAAATGGT ATATTTGCAA ACGAAAAATA ATTTGTAAAT AAAACTTTTA
2951  TGTATGTATT CTTAACGATC TAGCAGCAAA GGCTGAAAAA TAAACTTCGA
3001  TGAAAAATCT CAAAATCAAC TCTTAAAATT TAAATTTTGG CTTATAAGTA
3051  TAGTTCCTAA CTAGTTTAGA AGAAAAAATA TTTAAAGCGG GGAAGAGGAA
3101  AAGGAATAAA CTAATAGCTA AATTATTGCA TGCATGTAGC GATTTGAGGA
3151  CGACCGAGTT GTTTTGTCTG GATCAGCCGA CCGAGACAGA GCAATCTTCT
3201  TTAATCATAA ATAACCAGAA AAACCATACC AGTTCATCAC AATGGACCGA
3251  GTCAGAGTCA TTACATATTT TTCATTGTTG CGCACAGGAT TCACCATGTT
3301  CTTATGGGAA ATATTTTTAA CTCTCAAATG GTTATGATTT TGAACTCTCA
3351  TTTTTGAGAG AGAATTAACA AGCGAGCGAG CAATCAGGCC AAAAAGGGAG
3401  AAAGAAAATT ATTTTTGTTA ATTTTTTTTT AAGGTAGGGT GGAGGAGTCA
3451  TTACATGATT TTTTTTTATA TTCCCTCGTT GATTATATGC TGTTCAAATG
3501  GTTATGATTT TTTTAAAAGA TAACAACAAT ACAAATTAGT ATGTGATAGA
3551  TCATTTCACG AGCATATAGG ATTAAATTTA ACTTCTGTAA ATTACAAAAC
3601  AAACAAGTTT AACTGTTAAT ATACATTAAA TTTGTTTTTT TCAACTTAGG
3651  AATTGAATTT TATGTATATA TTTGTAAAAT GATATATTAA TTTATTTTTT
3701  TAAAAAAATA ATTATTTAGA TAACACGCAA ACTAGAAAAC CACCGCAGAA
3751  GTTCTCATAT TTCTTGTCCT ATCTGCACTT GCAGAGCTGA TGCTGCTGGG
3801  CTTCATATCC CTGCTTCTCA CCGTGGCACA GGCGCCCATC TCCAAGATCT
```

FIGURE 8 cont'd

```
3851   GCATCCCCAA GTCGGCTGCC AACATCTTGT TGCCGTGCAA GGCAGGCCAA

3901   GATGCCATCG AAGAAAGAAG CAGCAAGTGG TCGCCGGTCC TTGGCCGGCG

3951   CCGGCGGCGG GGACTACTGC TCGAAATTCG ATGTGAGAAT AACACCAGCT

4001   GCCGGCAAGC ACAACCTCGA TGCAATAACT AATTTAACTA TAATTGATTT

4051   TTCTTGGGTT TTCTGCAGGG CAAGGTGGCG CTGATGTCGG CAAAGAGCAT

4101   GCACCAGCTG CACATTTTCA TCTTCGTGCT CGCCGTGTTC CATGTTACCT

4151   ACTGCATCAT CACCATGGGT TTAGGGCGCC TCAAAGTGAG TTTGTCGTTC

4201   TGTCCCTCAT GCACATGTTT TCTCTAGTTC TAGCAAGATT GTCAGTCCTT

4251   CAAATGGATT GTTTCGACAA GAAACCCAAT TTATTAATTT GCCAGTAAAT

4301   ATATAATAAT TGATCTTTCT TGGTTTTAGA TGAAGAAATG GAAGAAGTGG

4351   GAGTCACAGA CCAACTCATT GGAGTATCAG TTCGCAATCG GTAGTGAATT

4401   AAGAATCTCC CTAACTATTT CATTTCAGAA CCTTTATGAT AATGTCTTGA

4451   AAGAGGAGGA GCAAATCAGC TGAAAAATAT GATCGATCCA TGCAGATCCT

4501   TCACGATTCA GGTTCACGCA TCAGACGTCG TTCGTGAAGC GGCATCTGGG

4551   ATCATTCTCA AGCACCCCTG GGCTCAGATG GATCGTGAGT TATCAATCTC

4601   CGAATACATG CTTGTTTTTT ATTCTTGCAA CTGGCCTAGC TGTTCCAATT

4651   CAATCCATAT TTTTTGAAAA AAAAAATATT CATGCCGTGT TTGTTGTTAG

4701   GTAGCATTCT TCAGGCAGTT CTTTGGGTCC GTCACCAAGG TGGACTACCT

4751   GACCATGCGG CAAGGCTTCA TCAATGTATA TACTAATCAA ACCTGACCAA

4801   TTCAACATTG ATGATGCAAA CAGAGACCAG GTTTTTTTTT TCGAGTGTGC

4851   ATTGAGTAAT GGTTTTAGCT TCTTCTCTTT TGCAGGCGCA TTTGTCGCAG

4901   AATAGCAAGT TCGACTTCCA CAAATACATC AAGAGGTCTT TGGAGGACGA

4951   CTTCAAAGTT GTCGTTGGCA TCAGGTCCGT CCTCGCTTTA TTAATTATAG

5001   GACTCTTATA TTCAACATTT TTTTTATAAA GAAACATATT TAGTCTCCAG

5051   TTGTGTATGT GTATGTGGAT CTTGACACAT TTGGCTGGTT TTGCAGCCTC

5101   CCTCTGTGGT TCGTCGGAAT CCTTGTACTC TTCCTCGATA TCCACGGTAA

5151   TCCTTGTCCT ATTTCATTCT TTTTTTTACT CTCAAAACCT TGTTCTGAAT

5201   TGGTCTTATA ATCACCATCG ATTTTTTTTC AACTTTTTCC CCGCGTGTAG

5251   GTCTTGGCAC ACTTATTTGG ATCTCTTTTG TTCCTCTCAT CGTAAGAGCG

5301   AAATTTCCCT GTCCAAAGAA ACAGTTAACA TAATTAATTA TGCTTTAATT

5351   TATCATGAAA ATTAATATGA TCATATAACT AATGAACAAA CATTCATGTG

5401   AATGCCACCG TTGTCTCAGA TCGTCTTGTT AGTTGGGACC AAGCTAGAGA

5451   TGGTGATCAT GGAGATGGCC AAGAGATAC AGGACAGGGC CACTGTGATC

5501   CAGGGAGCAC CTATGGTTGA ACCAAGCAAC AAGTACTTCT GGTTCAACCG

5551   CCCTGACTGG GTCTTGTTCT TCATACACCT GACACTCTTC CATGTACATG

5601   TTTAAAACCT AAACCTTGCT GCTCAACTAC AAATAGTACT TTATCTTTCA

5651   CAATTAACAC CTAATTAACT AACATAGCAT CCATCCATTT GTGGCTACTG

5701   ATCGATGGGA CGACGGATCG ATCATCACCA GAACGCATTT CAGATGGCGC

5751   ATTTCGTATG GACTATGGTG TGTATGCTAC TTGCTTAGTT GTTGCCATTA
```

Figure 8 cont'd

```
5801   TCAGTTCTTA AGCAAATTAA GTGTGATGCA TGCACTGACT AATGAGACAA
5851   AAAATGACAC AGCTTGTTCA TCGATCTGGT TGTTTGTGT GTGACAGGCA
5901   ACACCTGGTC TGAAGAAATG CTTCCATGAA AATATTTGGC TGAGCATCGT
5951   GGAAGTCATT GTGGGGATCT CTCTTCAGGT GCTATGCAGC TACATCACCT
6001   TCCCGCTCTA CGCGCTCGTC ACACAGGTGA ACAAGCCATT CACAAATTCT
6051   ATTAGCCGTT TCTTAATTGA TGACACTGTT AATTTTTAGA CACACGTTTT
6101   GACCATTTGT CTTATTAAAA ATATTTATGT AATTATCATT TGAGTTGTTT
6151   TATCACTAAA AGTACTTTTT AAATAATTTA TATTTTGCAT TTGTACAATT
6201   CTTTTAATAA GATAATGGTC AAACATGTGT CCAAAAGTTA ACAGCATCAT
6251   CTATTAAGAA AAGGAGGGGT TTTTTTTTTT TGGAATTTTG CAAAATTTGT
6301   TCAAAATCAG TCCAAAACCT TTTTTTTTTT CGAAATTTCA GTTTCACTAC
6351   CAGTCCCCAT AAAATGTCTT TTCTTTATTT CCACAAGATT GAACCCATGA
6401   GATGCCCTTT GTGTTGGTAT GTGTTTTGGC CATCACTTGC AGATGGGATC
6451   GAACATGAAG AAGACAATTT TCGAGGAGCA AACGATGAAG GCGCTGATGA
6501   ACTGGAGGAA GAAGGCGATG GAGAAGAAGA AGGTCCGGGA CGCCGACGCG
6551   TTCCTGGCGC AGATGAGCGT CGACTTCGCG ACGCCGGCGT CGAGCCGGTC
6601   CGCGTCGCCG GTGCACCTGC TGCAGGTCAC AGGGCGGGTC GGACGCCCGC
6651   CGAGCCCAAT CACGGTGGCC TCACCACCGG CACCGGAGGG GACATGTACC
6701   CGGTGCCGGC GGCGGCTGCG TCTCGCCAGC TGCTAGACGA CCCGCCGGAC
6751   AGGAGGTGGA TGGCATCCTC GTCGGCCGAC ATCGCCGATT CTGATTTTTC
6801   CTTCAGCGCA CAACGGTGAC GGGGGCGATC GGTTTCTGTA TTGATGCTGT
6851   ACCAAACATA GGAGTTTAAT ATATATATAA TTGTTACGGT AAAATCTAAT
6901   TATTGTGCGC GCACTTATAT TAGTCTTATA GCGCGACTGG TTCGTGATTA
6951   GACAAGGTGA TGCATGCTGT TTAGTTATAA AGGATATCAG CGCAGCTAAA
7001   AAAACTTACT CCCTACTTAA TAGATGACCT CGTTGATTTT TAACATTATT
7051   CGTCTTATTT AAAAAATTTA TGCAAATGTT TAAAACATAA ATCATGCTTA
7101   AAGTACTTTT AGTGATAAAA CAACTTACAA CAAAATAAAT TATACTTACC
7151   TAATTTTTTT TAATAAATCG AATGG
```

```
   1  TTATACCATG TGAGAAAGGC TGGAAGCATA TGCTCTTAGC AGGGACGCGT
  51  GCATGTTTAT ATAGGAGGCA TAAGCCGAAG AGATATACAT GAGGAGAGGT
 101  TTAAGATCAG TCTATCTTAT TTACAGTTTA AACACAAGGA GATAGAAAGA
 151  GATCCTAACC TACACATGTT ATACAAGTCA CGTATAATAC AAGAGTTATT
 201  TCGTCTAACA CCCTCCCCTC TGATATGATA AGTCGCCGGG AGAGAGAGAG
 251  AGTGTGTGGC TGCCCTCGCT GCACTGCACG CACATGTTTA CTTCTCCGAC
 301  TGAAACCACG GTGAAACCGG CGGCGGTGTC GCACTCCCCT GACTTTCCTC
 351  GCCGGGGTCC CGTCCGGACA ATTAAACCGT CTGTACCTGC CGGGCGTCGA
 401  CCCGATCGTG ATGTGGCGCC GCTTTGTCTG CAGCGAGCTG CGTGGCCGAT
 451  GGCAACAAAA CTGCGGTCAC ATACATGCAT ACCCCGCATA CCCCGACGCT
 501  CACCAGTAAG TAGGCTGTGG TGCGGCACCA CGGGCTCGCC GCCATTCATG
 551  CCATGCATGG GCCACCCGCC GGCGAAACCG CGGCGCTGCT GCCTGCCACC
 601  CCGCCGCCGT TGACGAAGAC TTCGCCCGGC CATCCATAAA AGCATGCATG
 651  GCTTGCTCTC ACCGGTCCGG CCACACACAC CACACTTCAC TTCGCCATTC
 701  GCACCACCGA GAGCGTAGCG TAACGTGTGT TTGAAGTCCT ACCATTAATT
 751  TTGCTGGATC GATGGCTGGG CCGGCGGGAG GTCGGGAGCT GTCGGACACG
 801  CCGACGTGGG CGGTGGCGGT AGTCTGCGCC GTCATGATAC TCGTCTCCGT
 851  CGCCATGGAG CACGCGCTCC ACAAGCTCGG CCACGTACGT GCTCTCGGTT
 901  CACTAGTGCT TAACTGTTTT TGATGTTTTC GGGCGTGTTT GGTAGCCTGC
 951  ATGGAGAGTG TATGAGCCCA AAAGTTCCCT CCCCGACCCA CTTTTCGCTG
1001  TTTGGTAGGG TGTATGGGCT GAGGAGAGCA TGCATCAACT GATGCAAAAA
1051  GGGCCTCAGC ATAGCTGAGC CCAGCACCCC CGCAGAGGCG AGCTGAGGCG
1101  AGTTATGCTG AGCCCATGCA CCCTCGCCCC GTCGCCCCGT CGCCCCGTCG
1151  CTCCCCCCCT GCACCTCTTC CTCCTCCCTC TTCCTACCAA ACACAGTCTC
1201  ATCCAAACAT GTAACAACAC ATGCATGACC ACCAAACAAC TGAAGATGAA
1251  TGTATTCATC ATGTCTATAC TTACCATGCA TCAACAGGGA ACAACTATGC
1301  TAGGGTGAGA ACAGCTGCCA AACACACCCG TGCACCTACT CATGCTGTGC
1351  CGGCGCTGGC GTACGTGTGC AGTGGTTCCA CAAGTGGCGC AAGAAGGCCC
1401  TGGGGGAGGC GCTGGAGAAG ATGAAGGCGG AGCTCATGCT GGTGGGCTTC
1451  ATATCCCTGC TCCTCATCGT CACGCAGGAT CCCGTCTCCA GGATCTGCAT
1501  CTCCAAGGAG GCCGGCGAGA AGATGCTCCC GTGCAAGCCT TACGACGGCG
1551  CCGGCGGTGG CAAAGGCAAG GACAATCACC GGAGGCTTCT CTGGCTCCAA
1601  GGCGAGAGCG AGACCCACCG CCGGTTCCTG GCTGCCCCGG CCGGAGTGGA
1651  CGTCTGCGCC AAACAGGTGA GCACCTAGCG TCGCCACAAA CCACAAACTA
1701  GCTAATGAGC ATGGACCTGA ATTTCTTCTC TTCTTGGCTT GGCTTGACTA
1751  AATTGGTTGT GCAGGGCAAG GTGGCGCTGA TGTCAGCGGG AAGCATGCAC
1801  CAACTGCACA TATTCATCTT CGTGCTCGCC GTCTTCCACG TCTTGTACAG
1851  CGTCGTCACC ATGACCCTAA GCCGTCTCAA AGTGAGCATC ATACTCGAGC
```

```
1901 TGTTTGTCAA TAATCCTTGG TTTCCAATCC AATTCCAAAG CTGGCACTGA
1951 TCCTGCTCCG GCTTCCTGCA GATGAAGCAA TGGAAGAAGT GGGAGTCGGA
2001 GACCGCCTCG CTGGAGTATC AGTTCGCGAA TGGTCAGCTT CAACTTTTCT
2051 TACTGAAACC GGATGCATTT ACAACAAACG CACGCACGAT CAATCATCAC
2101 AGTGTGAGCC GATACGTTGA ACCGATTGAA TCCTCGCAGA TCCATCGCGG
2151 TGCCGGTTCA CGCACCAGAC GACGTTGGTG AGGCGGCACC TGGGCCTCTC
2201 CAGCACCCCC GGCGTCAGAT GGGTGGTGGC CTTCTTCAGG CAGTTCTTCA
2251 CGTCGGTGAC CAAGGTGGAC TACCTGACCT TGCGGCAGGG CTTCATCAAC
2301 GCGCATCTCT CGCAGGGCAA CAGGTTCGAC TTCCACAAGT ACATCAAGAG
2351 GTCGTTGGAG GACGACTTCA AAGTCGTCGT CCGCATCAGG TACGCGCCAT
2401 TCCTTTCTCT GCACAAATTA ATACATCCAC CACCACATAG GTAGATAGAT
2451 AGATCGATAG ATAGATTATA CAAGTGCCGG TACGTACGTA CGTCTCATAT
2501 GATCTTGACA CATCTGTCCT CTTGCCGCAG TCTCAAGCTC TGGTTCGTGG
2551 CGGTCCTCAT CCTCTTCCTT GATTTCGACG GTAGCCGCCT TGTCCATGCC
2601 CTGCTCGCCC TCTCCTCCGC TTCTCTCCAT AATTTGTGAA CTTGTCCCGT
2651 ATATAACCAC ACCACCGTCG TCTTCTCGCA GGGATCGGCA CTCTTCTCTG
2701 GATGTCCGTG GTTCCTCTCG TGGTAAGTCC ACAATTTGAA TAGACAACCT
2751 GTCCAATTGT GATGTACAGT ACCTCCAAAC TTAATTAACA TGTCATTTGC
2801 TGATGTCTTG CGTGTAACAT TAGATCCTCT TGTGGGTTGG GACCAAGCTG
2851 GAGATGGTGA TCATGGAGAT GGCCCAGGAG ATCCATGACC GGGAGAGCGT
2901 CGTCAAGGGT GCTCCCGCCG TCGAGCCCAG CAACAAGTAC TTCTGGTTCA
2951 ACCGGCCTGA CTGGGTCCTC TTCCTCATGC ACCTCACACT CTTCCAGAAC
3001 GCGTTTCAGA TGGCTCATTT CGTGTGGACA GTGGTACGTA CAAGTACTTG
3051 TCACTTCACT TAGGCTAACT CCAACAAACG ACCCCAAATT AATGGTCCGT
3101 CGCGTCTGTT TGGGGTATGT TTGGGGTAAA CGGACACAAA ACTCAATCCA
3151 ACGCGCGGTA GCAAACGAAC GTTTTTCCGT ACGTTTTCGT CCGCTTTCGC
3201 CCCATCCCAG CCCAAATTCG TTGACGTTGT TGCATCGCAG GCCACGCCCG
3251 GCTTGAAGAA ATGCTACCAC GAGAAAATGG CAATGAGCAT CGCCAAGGTC
3301 GTGCTGGGGG TAGCCGCCCA GATCTTGTGC AGCTACATCA CCTTCCCGCT
3351 CTACGCGCTC GTCACGCAGA TGGGCTCACA CATGAAGAGA AGCATCTTCG
3401 ACGAGCAGAC GGCCAAGGCG CTGACCAACT GGCGAAAGAT GGCCAAGGAG
3451 AAGAAGAAGG CCCGAGACGC GGCCATGCTG ATGGCGCAGA TGGGCGGCGG
3501 CGCGACGCCG AGCGTCGGCT CGTCGCCGGT GCACCTGCTC CACAAGGCCG
3551 GGGCGCGGTC CGACGACCCC CAGAGCGTGC CGGCGTCCCC GAGGGCCGAG
3601 AAGGAAGGCG GCGGCGTGCA GCATCCGGCG CGCAAGGTAC CTCCTTGTGA
3651 CGGGTGGAGG TCGGCCTCGT CGCCGGCGCT CGACGCTCAC ATCCCCGGTG
3701 CAGATTTTGG CTTCAGCACG CAACGTTGAC CGATCAGACA AGTTCCTTTT
3751 TTTTTCGGTG AATAGAAGCG TATCATTTCA TTGATAGACA GTAGAAATTA
3801 CAGGAATGGC TGTCCTACTA CTATGTACAC AAGGGCACAG CAAAGGATCA
```

Figure 9 cont'd

```
3851   TTGATCTTGT  TACAAGAGCA  GTAGAAAGGG  ATTGCTCTCC  ATTGATCTTG
3901   TTAAGTTGTA  TGTCACAAAT  TGTTGCAGAA  AAAAGTGTAT  GTCATCCCAA
3951   CCAAGAGCTG  AGTTTGTGAT  GATTCGTGCA  ATAAGAATTG  CAAGTTTCAC
4001   CGAGTCAAAA  ATGAAGCTTC  TAAGTACGCA  CCAACCAACG  GACTCTTTCA
4051   TCTCAACAAA  AGAACTGTAA  ATGGCAATAA  TTCTGATAAC  ATCGGAAGGG
4101   AGCTC
```

Figure 10

```
   1  ATGGCAGGTG GGAGATCGGG ATCGCGGGAG TTGCCGGAGA CGCCGACGTG
  51  GGCGGTGGCC GTCGTCTGCG CCGTCCTCGT GCTCGTCTCC GCCGCCATGG
 101  AGCACGGCCT CCACAACCTC AGCCATAAAA CCACCGCAGA AGTTCTCATA
 151  TTTCTTGTCC TATCTCCACT TGCAGAGCTG ATGCTGCTGG GCTTCATATC
 201  CCTGCTTCTC ACCGTGGCAC AGGCGCCCAT CTCCAAGATC TGCATCCCCA
 251  AGTCGGCTGC CAACATCTTG TTGCCGTGCA AGGCAGGCCA AGATGCCATC
 301  GAAGAAGAAG CAGCAAGTGG TCGCCGGTCC TTGGCCGGCG CCGGCGGCGG
 351  GGACTACTGC TCGAAATTCG ATGGCAAGGT GGCGCTGATG TCGGCAAAGA
 401  GCATGCACCA GCTGCACATT TTCATCTTCG TGCTCGCCGT GTTCCATGTT
 451  ACCTACTGCA TCATCACCAT GGGTTTAGGG CGCCTCAAAA TGAAGAAATG
 501  GAAGAAGTGG GAGTCACAGA CCAACTCATT GGAGTATCAG TTCGCAATCG
 551  ATCCTTCACG ATTCAGGTTC ACGCATCAGA CGTCGTTCGT GAAGCGGCAT
 601  CTGGGATCAT TCTCAAGCAC CCCTGGGCTC AGATGGATCG TAGCATTCTT
 651  CAGGCAGTTC TTTGGGTCCG TCACCAAGGT GGACTACCTG ACCATGCGGC
 701  AAGGCTTCAT CAATGCGCAT TTGTCGCAGA ATAGCAAGTT CGACTTCCAC
 751  AAATACATCA AGAGGTCTTT GGAGGACGAC TTCAAAGTTG TCGTTGGCAT
 801  CAGCCTCCCT CTGTGGTTCG TCGGAATCCT TGTACTCTTC CTCGATATCC
 851  ACGGTCTTGG CACACTTATT TGGATCTCTT TTGTTCCTCT CATCATCGTC
 901  TTGTTAGTTG GGACCAAGCT ACAGATGGTG ATCATGGAGA TGGCCCAAGA
 951  GATACAGGAC AGGGCCACTG TGATCCAGGG AGCACCTATG GTTGAACCAA
1001  GCAACAAGTA CTTCTGGTTC AACCGCCCTG ACTGGGTCTT GTTTTTCATA
1051  CACCTGACAC TCTTCCATAA CGCATTTCAG ATGGCGCATT TCGTATGGAC
1101  TATGGCAACA CCTGGTCTGA AGAAATGCTT CCATCAAAAT ATTTGGCTGA
1151  GCATCGTGGA AGTCATTGTG GGGATCTCTC TTCAGGTGCT ATGCAGCTAC
1201  ATCACCTTCC CGCTCTACGC GCTCGTCACA CAGATGGGAT CGAACATGAA
1251  GAAGACAATT TTCGAGGAGC AAACGATGAA GGCGCTGATG AACTGGAGGA
1301  AGAAGGCGAT GGAGAAGAAG AAGGTCCGGG ACGCCGACGC GTTCCTGGCG
1351  CAGATGAGCG TCGACTTCGC GACGCCGGCG TCGAGCCGGT CCGCGTCGCC
1401  GGTGCACCTG CTGCAGGTCA CAGGGCGGGT CGGACGCCCG CCGAGCCCAA
1451  TCACGGTGGC CTCACCACCG GCACCGGAGG AGGACATGTA CCCGGTGCCG
1501  GCGGCGGCTG CGTCTCGCCA GCTGCTAGAC GACCCGCCGG ACAGGAGGTG
1551  GATGGCATCC TCGTCGGCCG ACATCGCCGA TTCTGATTTT TCCTTCAGCG
1601  CACAACGGTG A
```

Figure 11

```
   1 ATGGCTGGGC CGGCGGGAGG TCGGGAGCTG TCGGACACGC CGACGTGGGC
  51 GGTGGCGGTA GTCTGCGCCG TCATGATACT CGTCTCCGTC GCCATGGAGC
 101 ACGCGCTCCA CAAGCTCGGC CACTGGTTCC ACAAGTGGCG CAAGAAGGCC
 151 CTGGGGAGG CGCTGGAGAA GATGAAGGCG GAGCTCATGC TGGTGGGCTT
 201 CATATCCCTG CTCCTCATCG TCACGCAGGA TCCCGTCTCC AGGATCTGCA
 251 TCTCCAAGGA GGCCGGCGAG AAGATGCTCC CGTGCAAGCC TTACGACGGC
 301 GCCGGCGGTG GCAAAGGCAA GGACAATCAC CGGAGGCTTC TCTGGCTCCA
 351 AGGCGAGAGC GAGACCCACC GCCGGTTCCT GGCTGCCCCG GCCGGAGTGG
 401 ACGTCTGCGC CAAACAGGGC AAGGTGGCGC TGATGTCAGC GGGAAGCATG
 451 CACCAACTGC ACATATTCAT CTTCGTGCTC GCCGTCTTCC ACGTCTTGTA
 501 CAGCGTCGTC ACCATGACCC TAAGCCGTCT CAAAATGAAG CAATGGAAGA
 551 AGTGGGAGTC GGAGACCGCC TCGCTGGAGT ATCAGTTCGC GAATGATCCA
 601 TCGCGGTGCC GGTTCACGCA CCAGACGACG TTGGTGAGGC GGCACCTGGG
 651 CCTCTCCAGC ACCCCCGGCG TCAGATGGGT GGTGGCCTTC TTCAGGCAGT
 701 TCTTCACGTC GGTGACCAAG GTGGACTACC TGACCTTGCG GCAGGCCTTC
 751 ATCAACGCGC ATCTCTCGCA GGGCAACAGG TTCGACTTCC ACAAGTACAT
 801 CAAGAGGTCG TTGGAGGACG ACTTCAAAGT CGTCGTCCGC ATCAGTCTCA
 851 AGCTCTGGTT CGTGGCGGTC CTCATCCTCT TCCTTGATTT CGACGGGATC
 901 GGCACTCTTC TCTGGATGTC CGTGGTTCCT CTCGTGATCC TCTTGTGGGT
 951 TGGGACCAAG CTGGAGATGG TGATCATGGA GATGGCCCAG GAGATCCATG
1001 ACCGGCAGAG CGTCGTCAAG GGTGCTCCCG CCGTCGAGCC CAGCAACAAG
1051 TACTTCTGGT TCAACCGGCC TGACTGGGTC CTCTTCCTCA TGCACCTCAC
1101 ACTCTTCCAG AACGCGTTTC AGATGGCTCA TTTCGTGTGG ACAGTGGCCA
1151 CGCCCGGCTT GAAGAAATGC TACCACGAGA AAATGGCAAT GAGCATCGCC
1201 AAGGTCGTGC TGGGGGTAGC CGCCCAGATC TTGTGCAGCT ACATCACCTT
1251 CCCGCTCTAC GCGCTCGTCA CGCAGATGGG CTCACACATG AAGAGAAGCA
1301 TCTTCGACGA GCAGACGGCC AAGGCGCTGA CCAACTGGCG AAAGATGGCC
1351 AAGGAGAAGA AGAAGGCCCG AGACGCGGCC ATGCTGATGG CGCAGATGGG
1401 CGGCCGCGCG ACGCCGAGCG TCGGCTCGTC GCCGGTGCAC CTGCTCCACA
1451 AGGCCGGGGC GCGGTCCGAC GACCCCCAGA GCGTGCCGGC GTCCCCGAGG
1501 GCCGAGAAGG AAGGCGGCGG CGTGCAGCAT CCGGCGCGCA AGGTACCTCC
1551 TTGTGACGGG TGGAGGTCGG CCTCGTCGCC GGCGCTCGAC GCTCACATCC
1601 CCGCTGCAGA TTTTGGCTTC AGCACGCAAC GTTGA
```

```
1     GTTGGTACAT AAAAGACTCT TCCTTTGTCT GTTTTTTGTT CCCAGATTCA
51    TCTTTACTTA TTGACTAAAT TCTCTCTGGT GTGAGAAGTA AAATGGGTCA
101   CGGAGGAGAA GGGATGTCGC TTGAATTCAC TCCGACGTGG GTCGTCGCCG
151   GAGTTTGTAC GGTCATCGTC GCGATTTCAC TGGCGGTGGA GCGTTTGCTT
201   CACTATTTCG GTACTGTTCT TAAGAAGAAG AAGCAAAAAC CCCTTTACGA
251   AGCCCTTCAA AAGGTTAAAG AAGAGCTGAT GTTGTTAGGG TTTATATCGC
301   TGTTACTGAC GGTATTCCAA GGGCTCATTT CCAAATTCTG TGTGAAAGAA
351   AATGTGCTTA TGCATATGCT TCCATGTTCT CTCGATTCAA GACGAGAAGC
401   TGGGGCAAGT GAACATAAAA ACGTTACAGC AAAAGAACAT TTTCAGACTT
451   TTTTACCTAT TGTTGGAACC ACTAGGCGTC TACTTGCTGA ACATGCTGCT
501   GTGCAAGTTG GTTACTGTAG CGAAAAGGGT AAAGTACCAT TGCTTTCGCT
551   TGAGGCATTG CACCATCTAC ATATTTTCAT CTTCGTCCTC GCCATATCCC
601   ATGTGACATT CTGTGTCCTT ACCGTGATTT TTGGAAGCAC AAGGATTCAC
651   CAATGGAAGA AATGGGAGGA TTCGATCGCA GATGAGAAGT TTGACCCCGA
701   AACAGCTCTC AGGAAAAGAA GGGTCACTCA TGTACACAAC CATGCTTTTA
751   TTAAAGAGCA TTTTCTTGGT ATTGGCAAAG ATTCAGTCAT CCTCGGATGG
801   ACGCAATCCT TTCTCAAGCA ATTCTATGAT TCTGTGACGA AATCAGATTA
851   CGTGACTTTA CGTCTTGGTT TCATTATGAC ACATTGTAAG GGAAACCCCA
901   AGCTTAATTT CCACAAGTAT ATGATGCGCG CTCTAGAGGA TGATTTCAAA
951   CAAGTTGTTG GTATTAGTTG GTATCTTTGG ATCTTTGTCG TCATCTTTTT
1001  GCTGCTAAAT GTTAACGGAT GGCACACATA TTTCTGGATA GCATTTATTC
1051  CCTTTGCTTT GCTTCTTGCT GTGGGAACAA AGTTGGAGCA TGTGATTGCA
1101  CAGTTAGCTC ATGAAGTTGC AGAGAAACAT GTAGCCATTG AAGGAGACTT
1151  AGTGGTGAAA CCCTCAGATG AGCATTCTG GTTCAGCAAA CCTCAAATTG
1201  TTCTCTACTT GATCCATTTT ATCCTCTTCC AGAATGCTTT TGAGATTGCG
1251  TTTTTCTTTT GGATTGGGT TACATACGGC TTCGACTCGT GCATTATGGG
1301  ACAGGTGAGA TACATTGTTC CAAGATTGGT TATCGGGGTC TTCATTCAAG
1351  TGCTTTGCAG TTACAGTACA CTGCCTCTTT ACGCCATCGT CTCACAGATG
1401  GGAAGTAGCT TCAAGAAAGC TATATTCGAG GAGAATGTGC AGGTTGGTCT
1451  TGTTNGTTGG GCACAGAAAG TGAAACAAAA GAGAGACCTA AAAGCTGCAG
1501  CTAGTAATGG AGACGAAGGA AGCTCTCAGG CTGGTCCTGG TCCTGATTCT
1551  GGTTCTGGTT CTGCTCCTGC TGCTGGTCCT GGTGCAGGTT TTGCAGGAAT
1601  TCAGCTCAGC AGAGTAACAA GAAACAACGC AGGGGACACA AACAATGAGA
1651  TTACACCTGA TCATAACAAC TGAGCAGAGA TATTATCTTT TCCATTTAGA
1701  GGATCATCAT CAGATTTTAG CTTCAAGGTC CGGTTTTGTG GTTTATACAT
1751  AAGTTATAGT GACTTGATTT TTTTGTTTTG TTACAAAGTT ACCATCTTTG
1801  GATTAGAATT GGGAAATTGA ATCTGTTGT ATATTGTATT ATTTGGAACA
1851  TTGTGGATGC CCATGGATAT GTTTCTGTTC
```

SUBSTITUTE SHEET (RULE 26)

FIGURE 16

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus

Hvmlo-H1
Mlo
Osmlo-Hlm
Atmlo-H1
Consensus